Galaxy Clusters: a Dynamical Perspective

Raffaella Capasso

München 2019

Galaxy Clusters: a Dynamical Perspective

Raffaella Capasso

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Raffaella Capasso
aus Neapel, Italien

München, den 27. Februar 2019

Erstgutachter: Prof. Dr. Joseph Mohr
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 4. April 2019

"If you thought that science was certain...
well, that is just an error on your part."

Richard P. Feynman

Contents

List of Figures

List of Tables

Zusammenfassung

Galaxienhaufen sind das Resultat von gravitativem Kollaps überdichter Regionen, die bei Prozessen im frühen Universum erzeugt werden. Anschließend wachsen sie durch Kollisionen und Verschmelzung mit anderen Haufen und Akkretion von Material aus der Umgebung. In diesem Szenario spielen auch baryonische Prozesse im Intracluster-Medium und in Galaxien eine Rolle, was Galaxienhaufen sowohl zu einer Testumgebung für Theorien zur Strukturbildung und Galaxienentwicklung als auch zu interessanten Objekten für die Kosmologie macht.

Die Verteilung von Galaxienhaufen in Masse und Rotverschiebung reagiert sehr empfindlich auf Änderungen kosmologischer Parameter wie der mittlere Materiedichte und der Amplitude der Fluktuationen im Universum. Doch um Daten von Galaxienhaufen voll auszuschöpfen, bedarf es einer verlässlichen Bestimmung ihrer Masse. Diese kann mit verschiedenen Methoden abgeschätzt werden (z.B. durch den Gravitationslinseneffekt, über Röntgenemission oder den Sunyaev-Zel'dovich Effekt), doch die Information in der Phasenraumverteilung gravitativ gebundener Systeme gewährt einen einzigartigen Blickwinkel auf die Eigenschaften dieser Objekte.

Diese Dissertation beschäftigt sich mit der Analyse der Phasenraumverteilung von Haufengalaxien auf der Basis der sphärisch symmetrischen Jeans-Gleichung. In Kapitel 1 werden die nötigen Grundlagen eingeführt, um die Entwicklung des Universums zu verstehen. Anschließend beschreiben wir die Theorie der Strukturbildung, die zum Kollaps der heute beobachteten Halos führt. Kapitel 2 beschreibt die Eigenschaften von Galaxienhaufen, die Methoden zu ihrer Beobachtung und ihre Bedeutung für die Kosmologie. Außerdem wird die Theorie hinter der Jeans-Analyse mit ihren Vor- und Nachteilen erklärt.

Im zweiten Teil dieser Arbeit stellen wir eine Reihe unserer wissenschaftlichen Veröffentlichungen zu diesen Themen vor. Kapitel 3 befasst sich mit einer Untersuchung von ~ 3000 passiven Haufengalaxien, die keine Emissionslinien zeigen, und aus einem Datensatz von 110 mit dem Sunyaev-Zel'dovich Effekt entdeckten Haufen ausgewählt wurden. Diese Galaxienhaufen wurden mit dem South Pole Telescope (SPT) beobachtet, und decken eine Rotverschiebung $0.2 < z < 1.3$ ab. Wir führen eine dynamische Analyse mithilfe der Jeans-Gleichung durch, bei der ein Navarro-Frenk-White Dichteprofil und fünf verschiedene Modelle für die Geschwindigkeitsanisotropie angenommen werden. Für diese führen wir

eine Bayessche Modellmittlung durch, um die verschiedenen Anisotropiemodelle zu kombinieren. Die Studie hat drei Ziele: (1) Eine Untersuchung der Eigenschaften der Orbits passiver Galaxien, (2) Hinweise auf dynamisches Gleichgewicht und Selbstähnlichkeit in der Pseudo-Phasenraumdichte zu finden, und (3) eine Bestimmung der Haufenmasse und ein Vergleich mit Massen, die durch unabhängige Kalibrierung zuvor veröffentlichter SPT Analysen gewonnen wurden.

Die zweite Veröffentlichung, die in Kapitel 4 vorgestellt wird, befasst sich mit der Kalibrierung der Relation zwischen Haufenmasse M und Anzahl der Tochtergalaxien λ für 428 Haufen im CODEX Katalog mit Rotverschiebungen bis zu $z \sim 0.7$. Der spektroskopische Datensatz wurde mit SPIDERS Beobachtungen gewonnen und beinhaltet ~ 7800 Haufengalaxien. Wir untersuchen die $\lambda - M_{200c} - z$ Relation, indem wir die Wahrscheinlichkeit extrahieren, mit der die beobachteten Geschwindigkeiten innerhalb jedes Haufens mit der modellierten (projizierten) Geschwindigkeitsverteilung im Phasenraum für ein gegebenes λ, Rotverschiebung z und abgeleiteter Masse M übereinstimmen. Diese Analyse liefert die derzeit beste Bestimmung des Rotverschiebungstrends in der $\lambda - M_{200c} - z$ Skalierungsrelation.

Schließlich führen wir in Kapitel 5 eine Kalibrierung der Relation zwischen Röntgenhelligkeit und Masse für 344 CODEX Haufen bis zu $z \sim 0.66$ durch, indem wir die Dynamik der Haufengalaxien nutzen. Spektroskopische Beobachtungen durch SPIDERS liefern hierfür insgesamt ~ 6600 rote Haufengalaxien. Der abgedeckte Bereich in Rotverschiebung und Masse ist komplementär zu zuvor veröffentlichten Ergebnissen, was uns eine kombinierte Analyse erlaubt. Im Gegensatz zur $\lambda - M$ Relation, müssen wir hier die Verunreinigung der Röntgenquellen durch zufällige Überlagerungen entlang der Sichtlinie berücksichtigen. Wir weisen jeder Quelle eine Wahrscheinlichkeit zu, eine zufällige Überlagerung zu sein, und erstellen damit einen Datensatz mit einer Verunreinigung von lediglich 5%. In beiden Arbeiten zur Relation zwischen Masse und Observablen berücksichtigen wir die Selektionsfunktion der Haufen, statistische Korrekturen aufgrund des Eddington- und Malmquist Biases, und zusätzliche systematische Effekte.

Wir fassen unsere Ergebnisse in Kapitel 6 zusammen und schließen mit einer Diskussion über zukünftige und laufende Projekte.

Abstract

Galaxy clusters are the result of the gravitational collapse of overdense regions that are seeded by processes in the early universe, followed by a sequence of mergers and accretion of surrounding material. In this scenario, baryonic processes associated with the intracluster medium (ICM) and the galaxies also play a role, making galaxy clusters important laboratories for investigations of structure formation and galaxy evolution as well as useful cosmological probes.

The distribution of clusters in mass and redshift is highly sensitive to key cosmological parameters, such as the matter density, and the amount of matter fluctuations in the Universe. To fully exploit galaxy cluster data, reliable cluster masses are needed. While clusters masses can be estimated from several observables (e.g. weak gravitational lensing, X-ray data, Sunyaev-Zel'dovich effect), the information residing in the phase-space distribution of gravitationally bound systems offers a unique view on the evolution and properties of these objects.

This thesis is dedicated to the dynamical analysis of the phase-space distribution of cluster galaxies, based on the spherically-symmetric Jeans equation. In Chapter 1 we introduce the basic framework needed to understand the evolution of our Universe. We will then describe the theory of structure formation, leading to the collapse of the halos we observe today. Chapter 2 is focused on the properties of galaxy clusters, the techniques employed to observe them, and their use as cosmological probes. We also present the theory behind the Jeans analysis, its advantages and drawbacks.

In the second part of this thesis, we present a series of original scientific studies we carried out on these topics. Chapter 3 is dedicated to the study of $\sim$3000 passive, non-emission line cluster galaxies drawn from 110 Sunyaev-Zel'dovich effect selected galaxy clusters. These clusters were observed within the SPT-SZ survey, and cover the redshift range $0.2 < z < 1.3$. We perform a dynamical analysis based on the Jeans equation, adopting a Navarro-Frenk-White mass profile and five different velocity dispersion anisotropy profiles. We then perform Bayesian model averaging to combine the results from the different anisotropy models. This study has three main goals: (1) study the orbital characteristics of the passive galaxies, (2) investigate evidence for dynamical equilibrium and self-similarity with the pseudo-phase-space density profile, and (3) constrain cluster masses and perform

comparisons with masses obtained through independent calibration in previously published SPT analyses.

The second study, presented in Chapter 4, focuses on the calibration of the richness–mass scaling relation of 428 galaxy clusters that are members of the CODEX sample with redshifts up to $z \sim 0.7$. The spectroscopic dataset we analyze was obtained in the SPIDERS program and contains ∼7800 red member galaxies. For each cluster, we study the $\lambda - M_{200c} - z$ relation by extracting the likelihood of consistency between the observed phase-space distribution and the modeled projected distribution for a cluster at that redshift and λ. This analysis currently provides the tightest constraints on the redshift trend of the $\lambda - M_{200c} - z$ scaling relation.

Finally, in a third study (Chapter 5), we perform the calibration of the X-ray luminosity–mass scaling relation on a sample of 344 CODEX clusters up to $z \sim 0.66$, using the dynamics of their member galaxies. Spectroscopic follow-up measurements have been obtained from the SPIDERS survey, leading to a sample of ∼6600 red member galaxies. The redshift and mass ranges covered by our sample are complementary to previously published results, allowing us to perform combined analyses. Contrary to the richness–mass scaling relation calibration, for this study we have to take into account the fact that X-ray sources are likely to be contaminated by random superpositions along the line of sight. We assign each source a probability of being a chance superposition, and produce a sample with a 5% contamination fraction. In both our mass–observable scaling relation studies, we carefully account for the effects of the cluster selection function, statistical corrections for the Eddington and Malmquist biases, and additional systematic effects.

We summarize our findings in Chapter 6, concluding with a discussion on related future and ongoing projects.

Part I

General Overview

In the first part of this thesis, we provide the reader with the basic ingredients needed to understand the relevance of the studies presented in Part II.

Chapter 1

Cosmological framework

Unveiling the nature of the Universe. This is the ambitious challenge that astrophysicists have undertaken. From small dust particles, to the largest collapsed structures, from planets to black holes, and everything in between. About 14 billion years of physical processes, of structures forming, evolving, colliding, collapsing. *Cosmology* is what ties all of this together.

Since prehistoric times, man has sought to make sense of his existence and that of the world around him in some kind of theoretical framework. However, cosmology did not really come of age as a science until the 20th century, with the advent of Einstein's theory of general relativity (GR, Einstein, 1916, see Section 1.1). Given the cosmological symmetries, this theory implies that the Universe should either be expanding or contracting. However, Einstein himself rejected this notion in favor of the accepted idea of a static Universe. It was not until 1929 that Hubble, after proving that the Universe was not made of just one galaxy, convinced the astronomical community that the Universe was actually expanding.

In spite of Einstein's "greatest blunder", his theories laid the foundations of modern cosmology. The next few decades saw considerable theoretical and observational developments. The discovery of the Cosmic Microwave Background in 1965 (CMB, see Section 1.2) provided evidence of an evolving Universe which was hotter and denser in the past.

Nowadays, scientists have converged on a standard cosmological model, known as ΛCDM, according to which the Universe is currently expanding at an accelerating rate, and $\approx 95\%$ of its content is constituted by dark energy and cold, non-baryonic dark matter, both of unknown nature. This model predicts that structure formation begins from the gravitational collapse of primordial quantum fluctuations, followed by a sequence of mergers and accretion of surrounding material. The distributions of these structures create a web-like pattern, the so-called *cosmic web* (Bond et al., 1996). This web consists of dense knots of clustered galaxies connected by filaments, with vast regions of cosmic voids in between. Galaxy clusters, laying at the intersections of these filaments, constitute the perfect laboratory for both astrophysical and cosmological studies.

This thesis is dedicated to a dynamical study of galaxy clusters. Our goal is to learn more about their mass distribution, deeply linked to cosmological properties of our Uni-

verse, while also gaining information on the formation and evolution of massive halos. In this chapter, we present an overview of the standard cosmological framework, with particular emphasis on the role played by galaxy clusters in cosmological studies.

1.1 The Homogeneous Universe

A fundamental pillar of modern cosmology is the assumption that, on sufficiently large scales (greater than a few hundreds Mpc), the Universe is both homogeneous and isotropic.

This idea was first employed implicitly by Einstein (1917), de Sitter (1917) and Friedmann (1922) as a simplifying assumption which allows to find exact solutions of Einstein's equations. Later, Milne (1935) explored the general relativistic models in more detail, and introduced this concept as the *Cosmological Principle*.

Over time, several attempts to justify the use of this assumption have been made, both scientific and philosophical. In fact, this principle, also indicates that there is no privileged position in space-time (*Copernican Principle*), without which the assumption of isotropy would not imply homogeneity. The most appropriate approach to this assumption is an empirical one: the Cosmological Principle is accepted because it agrees with observations. In particular, CMB data demonstrate that the level of anisotropy of the Universe on large scales is about one part in 10^5 (see Section 1.2).

In advocating the Cosmological Principle, Einstein was particularly inspired by the ideas of Ernst Mach. *Mach's Principle* (see, e.g., Raine, 1975) states that the physical laws are determined by the large-scale structure of the Universe. The Cosmological Principle achieves Einstein's idea of a simplified global structure of the Universe, enabling a similar simplicity in the local behavior of matter. At such large scales, the strongest of the four forces of nature (electromagnetic, strong, weak, gravitational) is gravity: this is then the only interaction we should be worrying about when formulating a physical description of the Universe.

General Relativity

Einstein's General Theory of Relativity (Einstein, 1916) sets the scene for a complete description of the theory of gravity on which the evolution of the Universe is grounded. The essence of his theory is to consider gravitation no longer as a force, but as space-time itself, the geometry of which is characterized by the metric tensor $g_{\mu\nu}$[1]. This tensor serves as a local ruler telling us about deformations of space and time. The metric tensor is related to the content of the Universe via the Einstein's field equations (Einstein, 1915)

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

where $G_{\mu\nu}$ is the *Einstein tensor*, G is Newton's constant, and c stands for the speed of light. From now on, we use natural units and set $\hbar = c = 1$. The geometrical properties of

[1] Greek indices, running from 0 to 3, span over the four space-time components, while the three space-components are denoted by latin indices.

space-time are encoded in $G_{\mu\nu}$. In fact, from the *Ricci tensor* $R_{\mu\nu}$ one can construct the *curvature scalar* R (Peebles, 1993, Eq. 8.83)

$$R \equiv g^{\mu\nu} R_{\mu\nu}. \tag{1.2}$$

On the other hand, the information on the matter distribution is enclosed in the energy–momentum tensor $T_{\mu\nu}$. The only tensor compatible with the symmetry assumptions of the Cosmological Principle is found for a perfect fluid, such that (Peebles, 1993, Eq. 10.49):

$$T^{\mu\nu} = (\rho + p)\, u^\mu u^\nu + p g^{\mu\nu}, \tag{1.3}$$

where p and ρ are the pressure and energy density of the fluid, while u^μ is the fluid's four-velocity with which it moves through space-time.

Friedmann-Lemaître-Robertson-Walker metric

To determine how the Universe evolves, we need to define the metric that solves the field equations. The interval between two events can be written as

$$\mathrm{d}s^2 = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu. \tag{1.4}$$

The metric tensor determines all the geometrical properties of space-time. If the distribution of matter is uniform, then the space is uniform and isotropic. This, in turn, means that one can define a universal time, or *proper time*, such that at any instant the three-dimensional spatial metric is identical in all places and in all directions. As there is no reason why time should pass at different rates at different locations, the temporal term simply becomes proportional to $\mathrm{d}t^2$. While the Cosmological Principle forbids a spatial dependence of the temporal term, it still allows for a temporal variation of the spatial one. This variation is represented by the time-dependent dimensionless *cosmic scale factor* (or *expansion parameter*) $a(t)$. It is more convenient to express our metric using a reference frame that expands in tandem with the expansion of the Universe. This *comoving* coordinate system thus factors out the effect of the expansion. We can now rewrite Eq. 1.4 in terms of the scale factor $a(t)$ and a 3-dimensional line element $\mathrm{d}l$

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}l^2, \tag{1.5}$$

which in comoving spherical polar coordinates (r, θ, ϕ) becomes

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2 \left[\frac{\mathrm{d}r^2}{1 - Kr^2} + r^2 \left(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 \right) \right]. \tag{1.6}$$

This is known as the *Friedmann-Lemaître-Robertson-Walker* metric (FLRW Friedmann, 1922; Lemaître, 1931; Robertson, 1935; Walker, 1937). Here, K is the *curvature parameter*, representing the curvature scale of the Universe. It can assume three values, each corresponding to different curvatures and geometries:

- $K = -1$; negative curvature - hyperbolic geometry, open Universe
- $K = 0$; no curvature - Euclidean geometry, flat Universe
- $K = +1$; positive curvature - spherical geometry, closed Universe

Friedmann equations

Inserting the FLRW metric (Eq. 1.6) into the field equations (Eq. 1.10) it is possible to derive two equations describing the dynamical evolution of the Universe, expressed by the time-evolution of the scale factor $a(t)$:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\left(\sum_i \rho_i + \rho_k\right), \tag{1.7}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i(\rho_i + 3p_i). \tag{1.8}$$

Dots represent time derivatives of the scale factor, $\dot{a} \equiv \mathrm{d}a/\mathrm{d}t$, $\ddot{a} \equiv \mathrm{d}\dot{a}/\mathrm{d}t$, and ρ_i and p_i represent the energy density and the pressure of the different cosmic components. In particular, $i = \mathrm{m}$ for non-relativistic matter density (dust, or more precisely baryons and cold dark matter), and $i = \mathrm{r}$ for radiation density (relativistic matter). The *curvature density* ρ_k is defined as

$$\rho_k = -\frac{3K}{8\pi G a^2}. \tag{1.9}$$

These equations can be solved once one has an equation of state relating the fluid's pressure and energy density, $p_i = w_i\rho_i$. It is worth noting that, for ordinary perfect fluids such as matter or radiation, the Friedmann equations as written in Eqs. 1.7 and 1.8 cannot have a static solution, as the second derivative of the scale factor is always negative.

The cosmological constant

The absence of static solutions of the Friedmann equations led Einstein to believe that the field equations of general relativity (Eq. 1.1) needed to be revised (Einstein, 1917). He therefore introduced an additional term, the *cosmological constant* Λ, which acted as a repulsive term compensating the attractive gravitational pull. This term was originally identified as a property of space-time itself, therefore the natural place to add it was in the Einstein tensor $G_{\mu\nu}$, as part of the geometry of space-time:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.10}$$

It is also possible to interpret Λ as an additional fluid present in the Universe, contributing to the energy-momentum tensor. In this interpretation, Λ constitutes a form[2] of *dark energy*. In this case, the field equations can be rewritten as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G\left[T_{\mu\nu} - \frac{\Lambda}{8\pi G}g_{\mu\nu}\right] = 8\pi G\hat{T}_{\mu\nu}. \tag{1.11}$$

[2]Other dark energy models replace it with a dynamical field.

Applying to $\hat{T}_{\mu\nu}$ the stress-energy tensor equation presented in Eq. 1.3, in the assumption of a perfect fluid, we recover the formal pressure and energy density associated with Λ:

$$p_\Lambda = -\frac{\Lambda}{8\pi G} \ , \quad \rho_\Lambda = \frac{\Lambda}{8\pi G} \ . \tag{1.12}$$

The introduction of the cosmological constant leads to a modification of the Friedmann equations 1.7 and 1.8:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\left(\sum_i \rho_i + \rho_k + \rho_\Lambda\right), \tag{1.13}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left[\sum_i (\rho_i + 3p_i) - 2\rho_\Lambda\right]. \tag{1.14}$$

These equations have now static solutions for a geometrically spherical, closed Universe ($K = +1$) and for non-negative values of ρ_i, p_i, and Λ. However, such static solutions are dynamically unstable. Small fluctuations will either cause it to start collapsing or start expanding (see Section 1.2.1).

Equations 1.12 also imply that the equation of state of the Λ-fluid is

$$w_\Lambda = \frac{p_\Lambda}{\rho_\Lambda} = -1. \tag{1.15}$$

After the discovery of the expansion of the Universe in the late 1920s there was no longer any reason to seek static solutions to the field equations. Nevertheless, the interest in Λ has decreased, being the subject of several studies on both conceptual and observational grounds. Recent analyses of the magnitude-redshift relation for Supernovae of Type Ia (Riess et al., 1998; Perlmutter et al., 1999) finally provided striking evidence for an accelerated expansion of the Universe, renewing the interest on the cosmological constant once again.

Expansion of the homogeneous Universe

The Friedmann equations can be recast into a single equation by differentiating Eq. 1.13 and inserting it into Eq. 1.14. The result is known as the *continuity equation*. For a single component having pressure p and energy density ρ, it reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}(a^3\rho) + p\frac{\mathrm{d}}{\mathrm{d}t}a^3 = 0. \tag{1.16}$$

This equation represents the local energy-momentum conservation. It is analogous to the first law of thermodynamics, $\mathrm{d}U + p\mathrm{d}V = 0$, expressing the change in internal energy of an expanding fluid. We can rewrite this equation as

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0, \tag{1.17}$$

which, using the equation of state $w = p/\rho$, becomes

$$\dot{\rho} + 3\rho\frac{\dot{a}}{a}(1+w) = 0. \tag{1.18}$$

This equation is solved by

$$\rho(t) = \rho_0 a^{-3(1+w)}(t). \tag{1.19}$$

Normal matter (pressureless, i.e. $w = 0$) is diluted as a^{-3} as expected, while the energy density of fluids with considerable pressure ($w > 0$), such as relativistic particles, is diluted by the expansion more strongly. This is due to the loss of momentum, which is a significant contribution to the total energy. Inserting the w_Λ from Eq. 1.15 we find that the energy density associated to the cosmological constant, ρ_Λ, remains constant during the expansion of the Universe, $\rho_\Lambda = const.$ The energy density of the Cosmological Constant has always been (and will always be) the same throughout the whole expansion history of the Universe. Its value can therefore be considered a fundamental constant of Nature.

Rewriting Eq. 1.14 in terms of w, for a single fluid, we get $\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho(1+3w)$. All fluids characterized by $w < -\frac{1}{3}$ accelerate the expansion, while the ones with larger pressure decelerate it. Furthermore, inserting Eq. 1.19 into the first Friedmann equation (Eq. 1.13) we can derive the evolution of the scale factor as a function of the cosmic time:

$$a(t) \propto t^{-3(1+w)/2}. \tag{1.20}$$

It is now useful to introduce the *Hubble parameter*

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}, \tag{1.21}$$

which describes the relative expansion rate of a homogeneous and isotropic FLRW Universe. At the present time, $t = t_0$, we follow the convention of setting the scale factor $a(t_0) = 1$. It is also convenient to express the Hubble constant H_0 in terms of the dimensionless parameter h

$$H_0 = h \cdot 100\,\mathrm{km\ s^{-1}\,Mpc^{-1}}. \tag{1.22}$$

Current measurements report a value of $h = 0.6774 \pm 0.0049$ (Planck Collaboration et al., 2016c).

The Friedmann equation 1.13 also relates the space-time curvature to the energy density. In fact, for a single fluid, it can be rewritten as

$$\frac{K}{a^2} = \frac{8\pi G}{3}\rho - \frac{\dot{a}^2}{a^2} = \frac{\dot{a}^2}{a^2}\left(\frac{\rho}{\rho_c} - 1\right), \tag{1.23}$$

where we introduced the *critical density* as

$$\rho_c(t) = \frac{3}{8\pi G}\frac{\dot{a}^2(t)}{a^2(t)} = \frac{3H^2(t)}{8\pi G}. \tag{1.24}$$

Let us consider a small test mass m sitting on the surface of an expanding sphere of radius $a(t)$, density ρ, and mass $M(a(t)) = 4/3\pi a(t)^3\rho$. The potential energy of the test mass, as seen by an observer at the center of the sphere, is $U = -G\dfrac{M(a(t))m}{a(t)}$, while its kinetic energy is $T = 1/2mv^2$. As we will see in 1.1, the expansion velocity is given by Hubble's law $v = Ha(t)$. The total energy of the test particle is then

$$E_{\rm tot} = T + U = \frac{1}{2}a(t)^2 m\left(H^2 - \frac{8}{3}\pi\rho G\right) = \frac{1}{2}a(t)^2 m\left(1 - \frac{\rho}{\rho_{\rm c}}\right). \tag{1.25}$$

If the density of a fluid in an expanding sphere reaches the value of the critical density, the gravitational potential of the sphere is equal to its kinetic energy, halting its expansion. The matter density ρ_m contained in the Universe thus influences the balance between its expansion rate and the counter action of gravity, which in turn determines the geometry of the Universe:

- $\rho_m > \rho_c$: positive curvature ($K > 0$, spherical geometry), the Universe is closed; the potential energy will be greater than the kinetic one, halting the expansion, and the Universe will eventually start collapsing on itself (*Big Crunch*);
- $\rho_m = \rho_c$: zero curvature ($K = 0$, Euclidean geometry), the Universe is flat, forever expanding, with a decreasing expansion rate (recent measurements suggest that our Universe is most likely flat, but expanding at an accelerated rate caused by Λ; see Section 2.4);
- $\rho_m < \rho_c$: negative curvature ($K = 0$, hyperbolic geometry), the Universe is open; the expansion will prevail over the collapse, continuing forever.

The energy content of the Universe can be expressed in terms of dimensionless density parameters

$$\Omega_m(t) = \frac{\rho_m(t)}{\rho_{\rm c}(t)}\,, \quad \Omega_r(t) = \frac{\rho_r(t)}{\rho_{\rm c}(t)}\,, \tag{1.26}$$

$$\Omega_\Lambda(t) = \frac{\rho_\Lambda(t)}{\rho_{\rm c}(t)} = \frac{\Lambda}{3H^2(t)}, \tag{1.27}$$

$$\Omega_K(t) = \frac{\rho_K(t)}{\rho_{\rm c}(t)} = -\frac{K}{a^2H^2}. \tag{1.28}$$

These definitions, together with Eq. 1.19, allow us to rewrite the first Friedmann equation (Eq. 1.13) in the form

$$\frac{H^2(t)}{H_0^2} = \Omega_{r,0}\,a^{-4}(t) + \Omega_{m,0}\,a^{-3}(t) + \Omega_{k,0}\,a^{-2}(t) + \Omega_{\Lambda,0} =: E^2(a), \tag{1.29}$$

where the subscript 0 denotes values at the present epoch, $t = t_0$, and we substituted the value of w for normal matter ($w = 0$) and radiation ($w = 1/3$). As we followed the

convention of normalizing $a(t_0) = 1$, all contributions on the right hand side of Eq. 1.29 add up to 1, i.e. $\Omega_{\rm tot} = \sum_i \Omega_i = 1$.

Equation 1.29 highlights how the relevance of each energy component is dependent on time: radiation prevails at earlier times (*radiation-dominated* epoch), with the Universe expanding as $a \propto t^{1/2}$. Afterwards comes a time of *equality* between matter and radiation (*equivalence time*, $\rho_r = \rho_m$, $t = t_{\rm eq}$), followed by a *matter-dominated* era, during which the scale factor of the Universe evolves as $a \propto t^{2/3}$. At late times ($t \sim t_0$), the constant value of ρ_Λ begins to dominate, starting the *dark energy-dominated* epoch, leading to the current accelerated expansion of the Universe (see Figures 1.1 and 1.2).

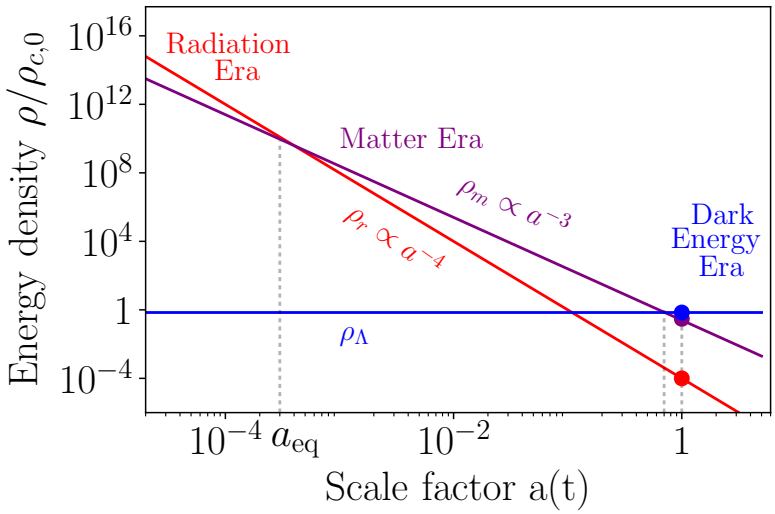

Figure 1.1: Energy density components of the Universe, in units of the critical density today, and their dependence on the scale factor $a(t)$. Radiation prevails at earlier times, with $a \propto t^{1/2}$. After a time of matter-radiation equality, matter is the dominant component, and $a \propto t^{2/3}$. At late times ($t \sim t_0$), the constant value of ρ_Λ begins to dominate.

Redshift and cosmological distances

The dawn of the observational cosmology era is marked by Hubble's discovery of the expansion of the Universe (Hubble, 1929). His observations of the galaxies outside the Local Group led to the finding of a correlation between their distance D to Earth and the velocity v with which they were receding from it

$$v = H_0 D. \tag{1.30}$$

This relation is known as the *Hubble law*. A consequence of an expanding Universe is the need to acquire different notions of *distance* between two points. Let us assume a photon with wavelength $\lambda_{\rm em}$ is emitted from a source at a time $t_{\rm em}$, and reaches the observer at $t_{\rm obs}$. While the photon travels, the scale factor of the Universe changes, increasing the distance between the two points. This causes a shift of the intrinsic wavelength towards

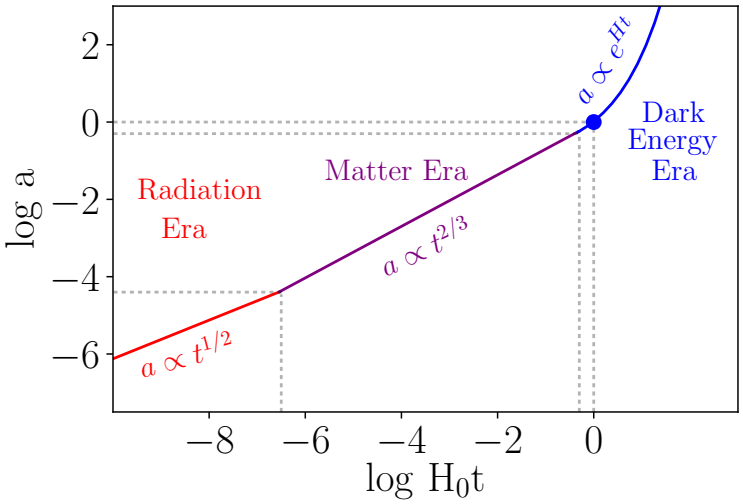

Figure 1.2: Evolution of the scale factor $a(t)$, as a function of cosmic time.

higher (redder) values, with $\lambda(t) \propto a(t)$. This increment defines a *(cosmological) redshift* z

$$z \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm em}} - 1 = \frac{a(t_{\rm obs})}{a(t_{\rm em})} - 1. \tag{1.31}$$

For a present time observer, $t = t_0$ and $a(t_{\rm obs}) = a(t_0) = 1$, which means that $a(t_{\rm em}) = 1/(1+z)$. The further away a galaxy is located from the observer, the smaller $a(t_{\rm em})$, the greater the redshift of that object.

The *proper distance* between us and an emitting source at redshift z, i.e. the distance measured by a ruler connecting the two points at a fixed time t, can be obtained from Eq. 1.5, setting $\mathrm{d}t = 0$. However, this distance is very unpractical, as it changes over time due to the expansion of the Universe. For this reason, it is convenient to introduce a *comoving distance*, in a static coordinate system, which is obtained by re-scaling the proper length of the differential photon path ($c\ \mathrm{d}t$) by $a(t)$:

$$\chi(z) = \int_{\chi(0)}^{\chi(z)} \mathrm{d}\chi = \int_{t_0}^{t} \frac{c\ \mathrm{d}t}{a(t)} = \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')}, \tag{1.32}$$

where the last equivalence comes from Eq. 1.29 and the fact that

$$\mathrm{d}t = \frac{\mathrm{d}t}{\mathrm{d}a}\mathrm{d}a = \frac{\mathrm{d}a}{H\,a} = -\frac{a\,\mathrm{d}z}{H(z)}. \tag{1.33}$$

For a flat Universe, this distance can be used to define the *angular diameter distance* D_A between two photons emitted at the same redshift, i.e. the ratio between the physical distance $\mathrm{d}\eta$ between the photons at the time of their emission and their differential angular separation $\mathrm{d}\theta$

$$D_A = a\,\chi = \frac{\mathrm{d}\eta}{\mathrm{d}\theta}. \tag{1.34}$$

This can also be interpreted as the proper distance between the source of the photons and the receiver, at the emission time. In a non-flat Universe, the comoving distance between the two photons, also known as *transverse* comoving distance, will instead be

$$\chi_T(z) = f_K(\chi(z)) \tag{1.35}$$

where $f(K)$ is a function of the curvature K, such that

$$f_K(\chi) = \begin{cases} \chi & \text{if } K = 0 \\ 1/K \sin(K\chi) & \text{if } K > 0 \\ 1/|K| \sinh(|K|\chi) & \text{if } K > 0 \end{cases} . \tag{1.36}$$

We can also define a *luminosity distance* D_L as the relationship between the bolometric observable flux F (i.e. the energy per unit time per unit area from the source to the observer) and the bolometric intrinsic luminosity L of the source:

$$D_L = \sqrt{\frac{L}{4\pi F}}. \tag{1.37}$$

This implies that the further the objects are from us, the dimmer they appear, making observations of objects at large redshift extremely difficult. However, observations of the apparent luminosity of the so-called *standard candles*, i.e. objects with known intrinsic luminosity, lead to the derivation of their luminosity distance. The luminosity distance is also linked to the angular diameter distance. Since

- the energy emitted by an object is reduced by the surface area $4\pi\chi^2 a_0^2$;
- each emitted photon loses energy as $E \propto a/a_0 = 1/(1+z)$;
- the rate at which photons are received per unit of time is diluted by a factor $a/a_0 = 1/(1+z)$;

we can write the *cosmological inverse squared law* as

$$F = \frac{L/(1+z)^2}{4\pi\chi^2 a_0^2}. \tag{1.38}$$

Inserting this into the definition of D_L (Eq. 1.37), and considering the definition of D_A (Eq. 1.34), we have that

$$D_L(z) = a_0(1+z)\chi = (1+z)^2 D_A(z). \tag{1.39}$$

This equation can be applied to standard candles (such as Supernovae Type Ia), with known redshift, to measure the Hubble parameter $H(z)$.

For completeness, we define the *look-back distance* D_T as the distance physically traveled by a photon in the *look-back time*. The look-back time is defined as the difference between the age of the Universe at the observation time (t_0) and the age of the Universe at the time the photons were emitted. From Eq. 1.33 we have that the physical photon path crossed in a small time interval $\mathrm{d}t$ is

$$\mathrm{d}D_T = c\,\mathrm{d}t = -\frac{c\,\mathrm{d}z}{(1+z)H(z)}, \tag{1.40}$$

which means

$$D_T(z) = \frac{c}{H_0}\int_0^z \frac{\mathrm{d}z'}{(1+z')E(z')}. \tag{1.41}$$

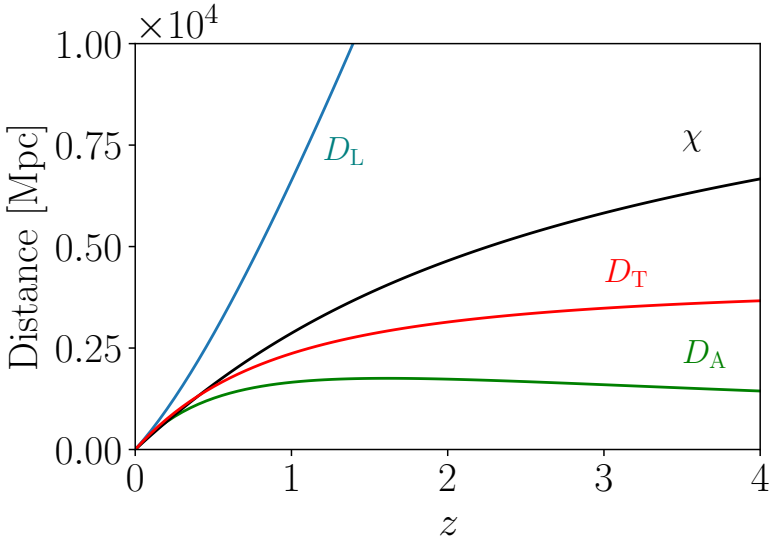

Figure 1.3: Comoving distance (black, Eq. 1.32), angular diameter distance (green, Eq. 1.34), luminosity distance (blue, Eq. 1.39), and look-back distance (red, Eq. 1.41) for an Universe with $\Omega_\mathrm{m} = 0.3$ and $\Omega_\Lambda = 0.7$. Note that D_A decreases at large redshift, due to the expansion history of the Universe. The rapid increase of D_L makes observations of objects at large z extremely difficult.

Figure 1.3 illustrates these definitions of distances as a function of redshift, for an Universe with $\Omega_\mathrm{m} = 0.3$ and $\Omega_\Lambda = 0.7$. We note how the angular diameter distance decreases at large redshift. This can be understood by considering that the observed shell of the Universe was smaller at higher redshifts.

1.2 Theory of structure formation

In Section 1.1 we have outlined the foundations of the standard cosmological model, starting from the assumption of an isotropic and homogeneous Universe. As already highlighted, this assumption is valid on large scales, greater than $\sim$ 100Mpc. On smaller scales, in fact,

the Universe is clearly clumpy, populated with several kinds of structures forming a rich web-like pattern. Long strands of galaxies form the filaments of this *cosmic web*, at whose intersections lie clumps of galaxies forming galaxy clusters, separated by large empty (or under-dense) regions. Section 1.2 is dedicated to the processes that led to the formation of such cosmic structures.

The discovery of Hubble's law, discussed in Section 1.1, led to development of two cosmological models: the Big Bang theory, advocated by Friedmann and Lemaître in the 1920s, and the steady-state theory, proposed by Hoyle (1948). The first model suggests that the Universe originated in a singularity with infinite density and temperature, while the second implies the continuous creation of matter to keep the density of the expanding Universe constant. For many years scientists argued over which theory was correct (Bondi & Gold, 1948; Hoyle & Narlikar, 1963; Kragh, 1996a,b), until observational evidence (e.g. the discovery of the CMB, Penzias & Wilson, 1965) began to support the idea that the Universe evolved from a hot dense state.

According to the Big Bang theory, the early Universe was extremely hot and dense. At this stage, ordinary matter was ionized, with electrons being free to wander and interact with photons. Radiation couldn't propagate for more than a very short distance before encountering an electron, making the Universe opaque. As the expansion on the Universe proceeded, the temperature decreased enough (≈ 3000K) to allow protons to capture electrons, and form neutral hydrogen atoms. This stage ($z \approx 1000$) is known as *recombination* or *decoupling* era. Light could propagate freely, and the Universe became optically thin. Today we can observe this relic radiation of the decoupling process, a snapshot of our Universe at the time of recombination: the Cosmic Microwave Background (CMB). The first observations of the CMB trace back to 1964, when Penzias & Wilson found a persistent isotropic radiation having a temperature of about 3.2 K. The CMB we see today is an isotropic radiation permeating the entire Universe, characterized by a black body spectrum. The expansion of the Universe has decreased the temperature of the CMB radiation by roughly a factor of 1000, down to an average temperature $T_{\rm CMB}$ of about 2.725 K. Despite having an extremely uniform temperature all over the Universe, this radiation still contains tiny temperature fluctuations, arising from random quantum fluctuations in the early Universe:

$$\frac{\Delta T}{T} = \frac{\Delta \rho_m}{\rho_m} \approx 10^{-5}. \tag{1.42}$$

Density and temperature fluctuations are tightly linked. The CMB photons retain memory of the matter and radiation distributions at the time of decoupling. If, at that time, a photon was in a slightly denser region, some of its energy had to be spent against the gravitational pull created by the overdensity, and therefore the region appears colder. On the other hand, photons passing through an underdense region spent less of their energy, making the region appear slightly hotter than average. These energy fluctuations reflect the pattern of the matter overdensities present in the early Universe, which are the seeds of the rich network of cosmic structure we observe today. Studies of the CMB allow us to gain insights into the dynamics and geometry of the Universe: its origin, evolution, and content.

Over the past couple of decades, many experiments have been dedicated to the study of the CMB, increasing more and more their accuracy (COBE, Smoot et al. (1992); WMAP, Komatsu et al. (2011); SPT, Carlstrom et al. (2011); ACT, Fowler et al. (2007); Planck, Planck Collaboration et al. (2014a), see Figure 1.4).

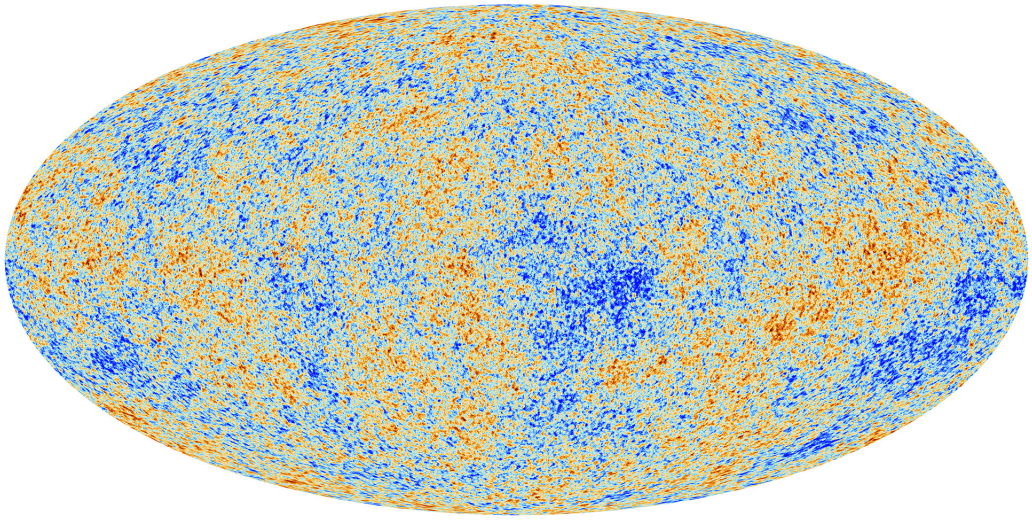

Figure 1.4: The anisotropies of the Cosmic Microwave Background (CMB) as observed by the Planck satellite. The tiny temperature fluctuations shown here ($\Delta T/T \sim 10^{-5}$) correspond to regions of slightly different densities, representing the initial seeds of all the structures we have in the Universe today. Credits: ESA, Planck Collaboration.

Cosmic inflation

The uniformity of $T_{\rm CMB}$ across the Universe raised the question of how photons emitted in widely separated regions could appear to be in thermal equilibrium at almost the same temperature, even though there was no time for these regions to interact before the photons were emitted. In fact, the finite speed of light sets a limit, known as *cosmological horizon*, to how far a photon can travel. Only photons closer than the horizon size can be in causal contact, interact, and exchange information. At the time of decoupling (t $\approx$ 300,000 yrs) the horizon size was $\approx$ 300,000 light years, which corresponds to about one degree in the sky today: two points on the surface of last scattering separated by an angle of more than a degree were out of causal contact at the time the CMB was emitted. This is known as the *horizon problem*.

The 1980s saw the emergence of a new model that could solve this issue: the *inflationary Universe* (Guth, 1981). The idea is that the Universe may have undergone a rapid period of inflation, from 10^{-37} to 10^{-32} seconds after the Big Bang, characterized by an accelerated exponential expansion which increased its size by a factor of $\approx 10^{26}$. The observable Universe then originated in a small causally-connected region. Two points that were initially in causal contact will be so quickly separated and causally disconnected that

there is no time for the homogeneity to be broken before the CMB is emitted. This theory provides an explanation to the homogeneity of the Universe on large scales, while also guaranteeing the existence of small fluctuations in the initial density field.

Jeans gravitational instability

The small fluctuations in the primordial density field led to the formation of the structures we see in the Universe today. Having provided a model that allows their presence in the early Universe, we have to define how they came to collapse and evolve.

In 1902, Jeans demonstrated the existence of an instability, known indeed as *Jeans gravitational instability*, which leads to the collapse of an accreting density fluctuation. He showed how small perturbations in the density of a fluid that is in average homogeneous and isotropic evolve with time. In particular, an overdense region is expected to accrete more and more material from its surroundings, thus becoming more dense, until it reaches a point of gravitational instability which can lead to the self-gravitational collapse of the fluctuation into a gravitationally bound object. A spherical inhomogeneity of radius λ and mass M, contained in a background fluid of mean density ρ, will grow if the self-gravitational force $F_{\rm g}$ is greater than the opposing pressure force $F_{\rm p}$

$$F_{\rm g} \simeq \frac{GM}{\lambda^2} \simeq \frac{G\rho\lambda^3}{\lambda^2} > F_{\rm p} \simeq \frac{p\lambda^2}{\rho\lambda^3} \simeq \frac{c_s^2}{\lambda}. \tag{1.43}$$

This fluctuation will thus grow with time if its lengthscale is greater than the *Jeans length* $\lambda_{\rm J}$ for that fluid

$$\lambda_{\rm J} = c_s \left(\frac{\pi}{G\,\rho}\right)^{1/2}, \tag{1.44}$$

where c_s is the speed of sound, and G the gravitational constant.
The same result is obtained by requiring the gravitational self-energy to be greater than the kinetic energy of the thermal motion of the gas, or by imposing the gravitational free-fall time to be shorter than the hydrodynamical time. Fluctuations having $\lambda > \lambda_{\rm J}$ will continue to grow, while those with $\lambda < \lambda_{\rm J}$ will propagate as acoustic waves with wavelength λ at velocity c_s. The same goes for a collisionless fluid, replacing the adiabatic sound speed with the mean square velocity v_* of the fluid particles. In this case, for $\lambda < \lambda_{\rm J}$ the self-gravity counteracts the tendency of particles to stream at the velocity v_*, while if $\lambda > \lambda_{\rm J}$ the particles undergo free streaming, and the fluid fluctuations are smeared out and dissipated.

The cosmic structure formation theory is mainly founded on this simple concept, generalized to include the effects of an expanding background and of a time-dependent matter density.

1.2.1 Linear growth of structures

To quantitatively describe the evolution of a density fluctuation ρ in a fluid with mean density $\bar{\rho}$, it is convenient to introduce the definition of *density contrast*:

$$\delta(\vec{x}, t) = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}. \tag{1.45}$$

To simplify our analysis, we make the following assumptions:

- DM and baryons evolve together and can be treated as a single fluid (this treatment effectively is valid after recombination);
- fluctuations accrete only via gravitational interactions, arising from an ordinary perfect fluid;
- perturbations are small ($\delta \ll 1$), implying that their evolution can be described in the linear regime, ignoring relativistic effects, and gravity treated in a Newtonian way.

Such a perfect fluid, in the Newtonian approximation, evolves following a set of equations linking its density ρ, velocity $\vec{v}$, pressure p and gravity Φ:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0; \quad \text{continuity equation (mass conservation)} \tag{1.46}$$

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla})\vec{v} + \frac{1}{\rho}\vec{\nabla} p + \vec{\nabla}\Phi = 0; \quad \text{Euler equation (momentum conservation)} \tag{1.47}$$

$$\nabla^2 \Phi - 4\pi G \rho = 0; \quad \text{Poisson equation} \tag{1.48}$$

where ∇^2 is the Laplace operator, and Φ the gravitational potential. Moreover, assuming the adiabatic condition, we have that the evolution in time of the entropy S is described by $\frac{\mathrm{d}S}{\mathrm{d}t} = 0$. This set of equations admits a static solution for $\rho = \text{const.} = \rho_0$, $\vec{v} = 0$, $p = \text{const.} = \mathrm{p}_0$, and $\Phi = \text{const.} = \Phi_0$. However, according to the Poisson equation, if $\rho_0 \neq 0$ the gravitational potential Φ cannot be constant. In other words, a homogeneous ρ distribution cannot be stationary. Therefore, the necessity of an expanding Universe can be found also in Newtonian, non-relativistic, gravity. This is also the reason why the static solutions of the Friedmann equations are unstable (see Section 1.1). Generalizing this approach to an expanding Universe leads to the correct results. For this reason we proceed with this analysis, with the simplifying assumption of Newtonian gravity.

We can now derive solutions for a small perturbation in the fluid by adding a small fluctuation to the static solution: $\rho = \rho_0 + \delta\rho$, $\vec{v} = \delta\vec{v}$, $p = p_0 + \delta p$, and $\Phi = \Phi_0 + \delta\Phi$.

Inserting these quantities into the set of equations 1.46-1.48 gives

$$\begin{aligned} &\frac{\partial \delta\rho}{\partial t} + \rho_0 \vec{\nabla}\cdot(\delta\vec{v}) = 0 \\ &\frac{\partial \delta\vec{v}}{\partial t} + \frac{c_s^2}{\rho_0}\vec{\nabla}\delta\rho + \vec{\nabla}\delta\Phi = 0, \\ &\nabla^2\delta\Phi - 4\pi G\delta\rho = 0 \end{aligned} \tag{1.49}$$

where we neglected terms of higher order in small quantities, and set the adiabatic speed of sound $c_s^2 = (\delta p/\delta\rho)_s$. For adiabatic density perturbations, time-dependent solutions of this set of equations can be found by solving the single differential equation

$$\frac{\mathrm{d}^2\delta\rho}{\mathrm{d}t^2} = 4\pi G\rho_0(\delta\rho) + c_s^2\nabla^2(\delta\rho), \tag{1.50}$$

which corresponds to a wave equation. Looking for solutions in the form of plane waves, we have

$$\delta\rho(\vec{x},t) \propto \delta(\vec{x},t)\exp(i\vec{k}\cdot\vec{x} + i\omega t), \tag{1.51}$$

with $\vec{k}$ and ω satisfying the *dispersion relation*

$$\omega^2 = c_s^2|\vec{k}|^2 - 4\pi G\rho_0 \equiv c_s^2(|\vec{k}|^2 - k_\mathrm{J}^2). \tag{1.52}$$

We introduced the *Jeans wave number* k_J, which is related to the Jeans length as

$$k_\mathrm{J} = \frac{\sqrt{4\pi G\rho_0}}{c_s} = \frac{2\pi}{\lambda_\mathrm{J}}. \tag{1.53}$$

If $\lambda > \lambda_\mathrm{J}$ the angular frequency ω is imaginary, and the solution will be a stationary wave of increasing or decreasing amplitude. On the other hand, if $\lambda < \lambda_\mathrm{J}$ the frequency is real and perturbations will oscillate as sound waves in the directions $\pm\vec{k}$.

The effect of expansion

The linear theory described above can be extended to the case of an expanding Universe. To do so, we need to introduce the scale factor into Eqs. 1.46-1.48. Namely, switching to comoving coordinates $\vec{r} = a\,\vec{x}$, the velocity of the fluid becomes $\vec{u} = \dot{\vec{r}} = H\vec{x} + \vec{v}$, and the fluid equations become:

$$\begin{aligned} &\frac{\partial\delta}{\partial t} + \frac{1}{a}\vec{\nabla}\cdot[(1+\delta)\vec{v}] = 0 \\ &\frac{\partial\vec{v}}{\partial t} + H\vec{v} + \frac{1}{a}(\vec{v}\cdot\vec{\nabla})\vec{v} + \frac{1}{a}\vec{\nabla}\Phi = 0, \\ &\nabla^2\Phi - 4\pi G\rho a^2\delta = 0 \end{aligned} \tag{1.54}$$

expressed in terms of the density contrast δ (Eq. 1.45). Assuming small perturbations, considering only linear terms in δ, and combining the three equations, we find that the time evolution of density perturbations can be expressed as

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\rho_0\delta + \frac{c_s^2}{a^2}\nabla^2\delta. \tag{1.55}$$

This represents a damped wave equation, with the left-hand side term $2H\dot{\delta}$, the *drag* term, acting against the growth of perturbations. As for the static case, the δ evolution depends on the interplay between gravity and pressure, and the *comoving* wave number

$$k_{\rm J} = \frac{\sqrt{4\pi G\rho_0}\ a}{c_s} \tag{1.56}$$

defines the threshold separating stable and oscillatory solutions. The general solutions to Eq. 1.55 can be expressed as

$$\delta(\vec{x}, t) = \delta_+(\vec{x})D_+(t) + \delta_-(\vec{x})D_-(t), \tag{1.57}$$

with D_+ and D_- representing the fluctuations' growing and decaying modes, respectively. From this equation we see that, in comoving coordinates, the spatial distribution of the fluctuations does not change with time: they are frozen. It is only the amplitude that changes.

One can assume that, after a short time, the decaying mode will become negligible and the perturbation remaining will just be described by the growing mode. This can be expressed as

$$D_+(z) \propto H(z)\int_z^\infty \frac{1+z'}{H^3(z')}dz'. \tag{1.58}$$

As an example, let us consider a flat ($K = 0$), matter dominated Universe ($\Omega_{\rm m} = 1$, $\Omega_\Lambda = 0$, Einstein-de Sitter). In this case (see Section 1.1), $a(t) \propto t^{2/3}$, $H(z) \propto (1+z)^{3/2}$, and we find the *growth factor* D_+ to be

$$D_+(z) \propto (1+z)^{-1} \propto a(t). \tag{1.59}$$

This therefore represents the way matter fluctuations evolved during the matter-dominated era of the Universe. Similarly, we find that before the equivalence time, during the radiation-dominated era,

$$D_+(z) \propto (1+z)^{-2} \propto a^2(t). \tag{1.60}$$

In the same way we can recover the growth of fluctuations for all other epochs of interest.

Adopting the normalization $D_+(t_0) = 1$ we would expect δ_+ to still represent the distribution of density fluctuations of the Universe today, provided the evolution remains linear. However, this is not the case as fluctuations grow with time, the condition $\delta \ll 1$ breaks down quickly, and their evolution becomes non-linear. In this regime, the complex

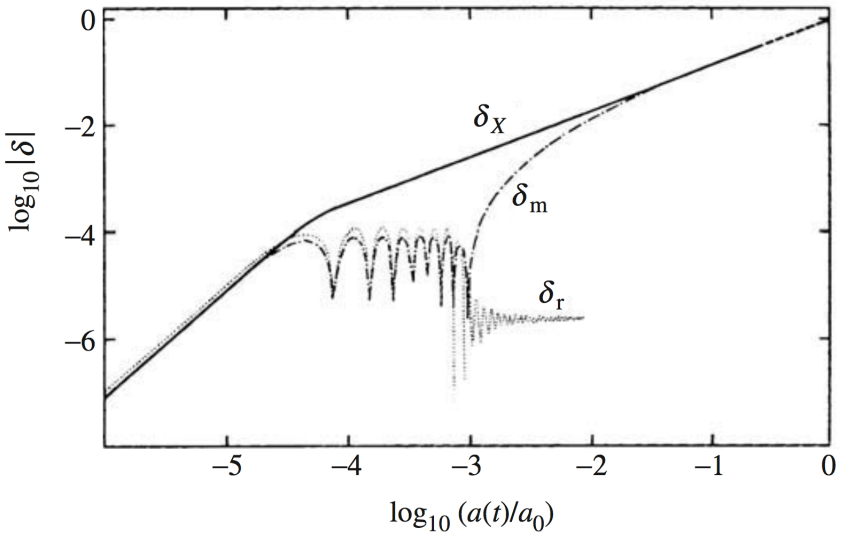

Figure 1.5: Evolution of density fluctuations for the baryonic matter δ_m, dark matter δ_X, and radiation δ_r components. These are shown here for a CDM Universe with $\Omega_{m,0} = 1$ and $h = 0.5$, at a mass scale of $10^{15} M_\odot$. Credits: Coles & Lucchin (2002).

evolution of structure formation can be studied only with numerical tools and N-body simulations, such as the Millennium Simulation (Springel et al., 2005).

Figure 1.5 illustrates the growth and evolution of overdensities corresponding to a mass scale $10^{15} M_\odot$, for the different components. Before the decoupling time, scales of the order of (or larger than) the horizon grow following the evolution of the main component. On smaller scales, baryonic and radiation fluctuations are smeared out by the radiation pressure. Oscillations in the baryon-photon plasma are produced, known as Baryonic Acoustic Oscillations (BAO). Inhomogeneities in the cold dark matter (CDM) component, on the other hand, being free from the effects of radiation pressure, start to grow due to gravitational interactions. After the decoupling time, the baryonic fluctuations can start to grow. In the meantime, dark matter overdensities had continued to accrete, forming compact structures with a potential well. The baryonic matter overdensity are then driven to collapse directly into the DM potential wells, causing structures to form faster than they would without the presence of DM.

Statistics of density perturbations

Describing the distribution of matter in an infinite space is not straightforward. One might try to divide the Universe into finite volumes, initially evolving independently of each other. However, gravitational forces between the cells would soon start to be strong, breaking this independence. Instead of considering a generic perturbation as a sum of spatial components, it is thus more useful to consider the perturbation as a superposition of plane waves, evolving independently while the fluctuations remain linear. To do so, we need to represent the matter distribution in Fourier space, in terms of the wavevectors

of each independent component $\vec{k}$. Let us consider a cube of volume V and side length $L \gg l_s$, where l_s is the maximum scale at which there is significant structure due to the perturbations. In this case, the cube is a *fair sample* of the Universe. A representative realization of the Universe can thus be constructed by dividing it into cells of volume V, with periodic boundary conditions at the faces of each cube. Denoting by $\langle\rho\rangle$ the mean density in the cell, and by $\rho(\vec{x})$ the density at a point specified by the position vector $\vec{x}$, the density contrast $\delta(\vec{x}) = [\rho(\vec{x}) - \langle\rho\rangle]/\langle\rho\rangle$ can be expressed as a Fourier series:

$$\delta(\vec{x}) = \sum_{\vec{k}} \delta_{\vec{k}} \exp(i\vec{k} \cdot \vec{x}). \tag{1.61}$$

The assumption of periodic boundary conditions at each surface of the cube requires the components of the wavevector $\vec{k}$ to be

$$k_j = n_j \frac{2\pi}{L} \; ; \;\; n_i \in \mathbb{Z} \; ; \;\; j = x, y, z. \tag{1.62}$$

The Fourier coefficients $\delta_{\vec{k}}$ are thus given by

$$\delta_{\vec{k}} = \frac{1}{V} \int_V \delta(\vec{x}) \exp(-i\vec{k} \cdot \vec{x}) \mathrm{d}\vec{x}. \tag{1.63}$$

The mean value of the perturbation $\delta_{\vec{x}}$ is equal to zero by definition, in all realization. However, there is a variance σ^2 in $|\delta_{\vec{k}}|$ between the different realizations:

$$\sigma^2 \equiv \langle\delta(\vec{x})^2\rangle = \sum_{\vec{k}} \langle|\delta_{\vec{k}}|^2\rangle = \frac{1}{V} \sum_{\vec{k}} \delta_k^2 \to \frac{1}{2\pi^2} \int_0^\infty P(k) k^2 dk, \tag{1.64}$$

where we took the limit $V \to \infty$. Also, we assumed that the density field is statistically homogeneous and isotropic, so that there is no dependence on the direction of k but only on $k = |k|$. The variance does not depend on the spatial position, but it does depend on the time evolution of δ_k. Therefore, σ^2 only provides us with information on the amplitude of perturbations, not on their spatial structure.

The quantity $P(k) \equiv \langle|\delta_{\vec{k}}|^2\rangle_{|\vec{k}|=k} = \delta_k^2$ is known as the *power spectrum*. If δ is described by a Gaussian random field, as it appears to be the case in our Universe (see, e.g. Planck Collaboration et al., 2016a), then different Fourier modes are chosen independently from a Gaussian distribution. The power spectrum then contains the full information about the field. $P(k)$ is also related to the *two-point correlation function* of density contrast,

$$\xi(r) = \langle\langle\delta(\vec{x})\delta(\vec{x} + \vec{r})\rangle_{\vec{x}}\rangle_{|\vec{r}|=r} = \frac{V}{(2\pi)^3} \int P(k) e^{-i\vec{k}\cdot\vec{r}} d\vec{k}, \tag{1.65}$$

where the average is taken over all locations $\vec{x}$ and all vectors $\vec{r}$ with length r. This function describes the probability of finding two overdensities separated by the distance $\vec{r}$.

Typically, the primordial power spectrum is assumed to be given by a power-law

$$P(k) = Ak^{n_s}, \tag{1.66}$$

with the exponent n_s usually referred to as *spectral index*. Inflationary models (e.g. Mukhanov & Chibisov, 1981) predict a value of n_s close to unity. Observationally, the most constraining results on the spectral index report a value of $n_s = 0.9649 \pm 0.0042$ (Planck Collaboration et al., 2018b).

The evolution of the primordial matter power spectrum is connected to the rate at which fluctuations grow on different scales. As illustrated in Section 1.2.1, this depends on when the fluctuations enter the horizon. Before the equivalence time, fluctuations on scales larger than the size of the horizon grow through self-gravity. On scales smaller than the horizon, the radiation density acts as a pressure that prevents the further collapse of any perturbation in the matter density. As the Universe expands, the size of the horizon becomes larger, encompassing density fluctuations on progressively larger scales, which stay frozen until the time of equivalence between matter and radiation. The last fluctuations to freeze are the ones having the same size as the horizon at the time of equivalence. At this scale $P(k)$ has a turn-over, and is characterized by its primordial shape, $P(k) \propto k$. The power spectrum we observe today is thus described by

$$P(k, z = 0) = T^2(k)Ak^{n_s}, \tag{1.67}$$

where T^2 is the *transfer function*, representing departures from the primordial shape. In particular, defining $k_{\rm H}$ as the horizon scale at matter-radiation equality, we have

$$T(k) = \begin{cases} 1, & k \ll k_{\rm H} \\ k^{-2}, & k \gg k_{\rm H} \end{cases},$$

with a turn-over at $k = k_{\rm H}$.

After the equivalence time, the growth of fluctuations follows Eq. 1.58. This linear growth influences δ independent of position, growing all Fourier modes δ_k equivalently, such that $P(k, z) = D_+^2(z)P(k, 0)$. As we will see in Section 1.2.2, the linear approximation greatly underestimates the growth of structures on scales small enough to collapse. The actual matter power spectrum is modified by a correction for non-linear evolution. Figure 1.6 shows the comparison between the linear and non-linear power spectrum. The transfer function suppresses the primordial $P(k) \propto k$ at scales smaller than ≈ 100 Mpc. The non-linear evolution starts being significant at scales smaller than ≈ 5 Mpc. At these scales we have to consider the non-linear growth and the effect of collapsed structures.

The amplitude of the power spectrum is often expressed in terms of σ_8, the variance of the matter field when averaged over a sphere of 8 h^{-1} Mpc radius. To compare data and theoretical results, as we actually observe the density field smoothed with some finite resolution, it is useful to consider the density fluctuations filtered on a resolution scale R:

$$\delta(\vec{x}; R) = \int \delta(\vec{x'})W(|\vec{x'} - \vec{x}|; R)\mathrm{d}^3x', \tag{1.68}$$

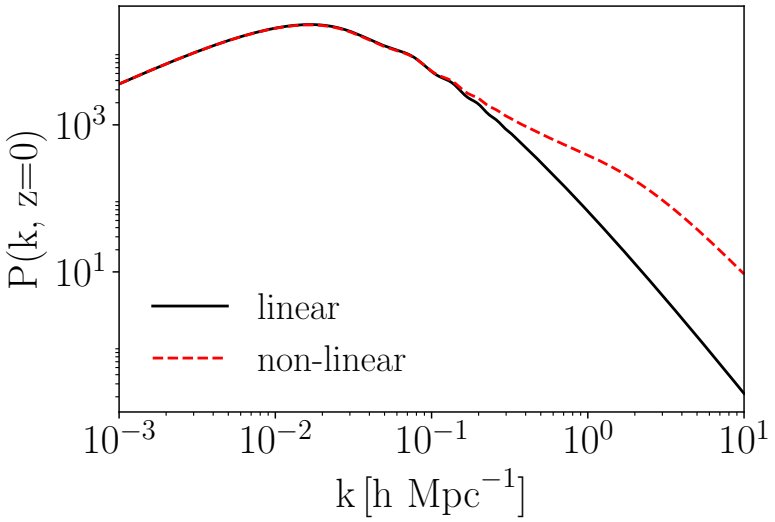

Figure 1.6: Linear (black) and non-linear (red) matter power spectrum at $z = 0$. We note how the non-linear evolution starts being significant at scales smaller than ≈ 5 Mpc.

where $W(\vec{x}; R)$ is the *window function* filtering the density field. One of the most commonly used functions is the top-hat filter, which in real-space is defined as

$$W(\vec{x}; R) = \begin{cases} \text{const.}, & |x| \leq R \\ 0, & \text{elsewhere} \end{cases},$$

while its Fourier transform is

$$W(kR) = \frac{3(\sin(kR) - kR\cos(kR))}{(kR)^3}. \tag{1.69}$$

The variance inside the volume contained in the sphere of radius R and mass M can be expressed as

$$\sigma_M^2 = \frac{1}{2\pi^2} \int_0^\infty P(k) W(kR)^2 k^2 \mathrm{d}k. \tag{1.70}$$

Assuming a scale-invariant power spectrum of the form $P(k) \propto k^{n_s}$, we get

$$\sigma_M^2 \propto \int_0^\infty k^{n_s+2} \mathrm{d}k \propto k^{n_s+3} \propto R^{-(n_s+3)} \propto M^{-(n_s+3)/3}. \tag{1.71}$$

This implies that, as long as $n_s > -3$, larger density fluctuations form smaller objects. This picture is referred to as the *hierarchical clustering scenario*.

1.2.2 Non-linear growth and spherical collapse

When the above approximation for $\delta \ll 1$ ceases to be valid, density fluctuations collapse, and dense, gravitationally bound structures form. As the total matter must be conserved,

the presence of overdense regions with $\delta \gg 1$ implies the existence of underdense regions, with $\delta < 0$. These are referred to as *voids*.

As mentioned above, an accurate treatment of the evolution of structures in the non-linear regime requires the use of numerical simulations. However, we can gain some insight by analyzing a simplified model: the case of a *spherical collapse*. Let us consider a sphere in an initial state i, immersed in an expanding flat Einstein-de Sitter (EdS, $\Omega = \Omega_{\rm m} = 1$) Universe. All properties of this sphere will be denoted by a $\sim$ superscript. The initial expansion of this EdS Universe is described by H_i and $\tilde{a}_i$, and the sphere is characterized by a constant overdensity $\tilde{\delta}_i$ with respect to the background Universe. We saw that for a matter-only, EdS, Universe

$$a(t) = a_i \left(\frac{t}{t_i}\right)^{2/3}. \tag{1.72}$$

Considering that $H_i = H(t_i) = \dfrac{\dot{a}_i}{a_i} = \dfrac{2}{3t_i}$, and that the mass conservation $\rho a^3 = \text{const.}$ implies $\rho t^2 = \text{const.}$, we find

$$\rho(t) = \frac{\rho_{\rm c,i} t_i^2}{t^2} = \frac{3H_i^2}{8\pi G}\frac{4}{9H_i^2}\frac{1}{t^2} = \frac{1}{6\pi G t^2}, \tag{1.73}$$

where $\rho_{\rm c,i}$ is the initial critical density. For the perturbation we have that

$$\tilde{\Omega}(t_i) = \frac{\tilde{\rho}(t_i)}{\rho_{\rm c}(t_i)} = \frac{\rho(t_i)(1+\tilde{\delta}_i)}{\rho_{\rm c}(t_i)} = \Omega(t_i)(1+\tilde{\delta}_i). \tag{1.74}$$

Thus, any perturbation with $\tilde{\Omega}(t_i) > 1$ will evolve as a closed Universe of its own, with $\tilde{\Omega}_i = 1 + \tilde{\delta} > 1$. The expansion of the perturbation is described by the Friedmann equation

$$\left(\frac{\dot{\tilde{a}}}{a_i}\right)^2 = H_i^2 \left[\tilde{\Omega}_i \frac{a_i}{\tilde{a}} + 1 - \tilde{\Omega}_i\right], \tag{1.75}$$

where we have multiplied both sides by $(\tilde{a}/a_i)^2$ (see Coles & Lucchin, 2002, Eq. 14.1.4). Since the metric of the overdense sphere is closed, the scale factor $\tilde{a}$ will reach a maximum value at some time $t_{\rm m}$. Afterwards, the radius of the perturbation will start to decrease, the spherical region ceases to expand with the background Universe and instead begins to collapse, forming a structure. $t_{\rm m}$ is known as *turn-around time*. This transition happens when $\dot{\tilde{a}} = 0$. From Eq. 1.75 we then have that

$$\frac{a_i}{\tilde{a}(t_{\rm m})} = \frac{\tilde{\Omega}_i - 1}{\tilde{\Omega}_i}. \tag{1.76}$$

Knowing that $\rho a^3 = \text{const.}$, and $\tilde{\rho}(t_i) = \rho(t_i)(1+\tilde{\delta}_i) = \rho(t_i)\tilde{\Omega}_i$, we have

$$\tilde{\rho}(t_{\rm m}) = \rho(t_i)\tilde{\Omega}_i = \rho_{\rm c}(t_i)\frac{a_i}{\tilde{a}(t_{\rm m})} = \rho_{\rm c}(t_i)\frac{(\tilde{\Omega}_i - 1)^3}{\tilde{\Omega}_i^2}. \tag{1.77}$$

From the cycloid solution to the Friedmann equation in a closed, matter-dominated Universe (see Coles & Lucchin, 2002, Eq. 2.14.9) we can find the turn-around time as

$$t_{\rm m} = \frac{\pi}{2H_i}\frac{\tilde{\Omega}_i}{(\tilde{\Omega}_i - 1)^{3/2}} = \frac{\pi}{2H_i}\left(\frac{\rho_{\rm c}(t_i)}{\tilde{\rho}(t_{\rm m})}\right)^{1/2} = \left(\frac{3\pi}{32G\tilde{\rho}(t_{\rm m})}\right)^{1/2}. \tag{1.78}$$

Inserting this equation into Eq. 1.73 we recover the non-linear overdensity of the sphere at the time of turn-around:

$$\tilde{\delta}_{\rm m} = \frac{\tilde{\rho}(t_{\rm m})}{\rho(t_{\rm m})} - 1 = \left(\frac{3\pi}{4}\right)^2 - 1 \approx 4.6. \tag{1.79}$$

As the sphere begins to shrink, its density increases until virial equilibrium is reached. The result is an extended system with radius $\tilde{R}_{\rm vir}$ and mass $\tilde{M}_{\rm vir}$. The virial theorem provides us with the relation between kinetic and potential energy of the sphere, $E_{\rm kin} = -1/2\, E_{\rm pot}$. From here we have that the total energy is

$$\tilde{E}_{\rm vir} = -\frac{1}{2}\frac{3G\tilde{M}}{5\tilde{R}_{\rm vir}}. \tag{1.80}$$

Assuming that in the collapsing phase there was no mass loss from the system due to effects connected with shocks, and no loss of energy by thermal radiation, the energy and mass of the virialized system have to be the same as they were at the turn-around time. The energy at $t_{\rm m}$ is purely potential, therefore for a homogeneous sphere of turn-around radius $\tilde{R}_{\rm m}$

$$\tilde{E}_{\rm m} = \tilde{E}_{\rm pot} = -\frac{3}{5}\frac{G\tilde{M}_{\rm m}}{\tilde{R}_{\rm m}}. \tag{1.81}$$

This implies that $\tilde{R}_{\rm vir} = \frac{1}{2}\tilde{R}_{\rm m}$, and $\tilde{\rho}_{t_{\rm vir}} = 8\tilde{\rho}_{t_{\rm m}}$. Assuming $t_{\rm vir} \approx 2t_{\rm m}$, and that the background Universe is still described by an EdS model, the ratio between the density of the perturbation and that of the background is

$$\tilde{\delta}_{\rm vir} = \frac{\tilde{\rho}(2t_{\rm m})}{\rho(2t_{\rm m})} - 1 = 2^2 \cdot 8 \cdot (\tilde{\delta}_{\rm m} + 1) - 1 \approx 178. \tag{1.82}$$

Therefore, in a matter-dominated Universe, virialization occurs at matter overdensities of about 200. This is the reason why it is common to refer to halo properties in terms of the sphere inside which the overdensity relative to the mean matter density is 200, such as the mass M_{200} inside the sphere of radius R_{200}.

As already mentioned, the linear theory provides only an approximated description of the growth of perturbations in this regime. In fact, extrapolating the linear growth to $t_{\rm vir}$ yields an overdensity of (see Coles & Lucchin, 2002, 14.1.13)

$$\tilde{\delta}_{\rm vir,lin} \approx 1.686 \tag{1.83}$$

instead of 178. However, this value of $\tilde{\delta}_{\rm vir,lin}$ remains a good approximation on large scales, and can be used to find an estimated *halo mass function* (see Section 2.4).

Chapter 2

Clusters of galaxies

At the massive end of collapsed objects, we find galaxy clusters. These structures form through a series of mergers and accretion of surrounding material, reaching final virial masses of 10^{14} - $10^{15} M_{\odot}$ and sizes of the order of a few Mpc. This chapter is dedicated to the description of their properties, the techniques used to measure their mass, and their role as cosmological probes.

2.1 Components of the cluster matter

Galaxy clusters, due to their large mass and volume, are representative samples of the cosmic matter content. While their total mass is dominated by dark matter ($\sim$ 80%), baryonic processes prominently determine the evolutionary physics and the observational appearance of clusters. Relativistic plasma, AGNs, gas cooling and heating, galaxies, shocks, relics – all these processes and constituents reflect the formation history and the dynamical properties of clusters of galaxies. The knowledge of the complex interplay between the different components of the cluster matter is therefore necessary to fully exploit clusters as cosmological and astrophysical probes.

2.1.1 Galaxies

Stars residing inside the cluster galaxies produce the photons that made the first observations of clusters possible. In optical wavelengths, galaxy clusters appear as localized overdensities of galaxies, with hundreds to thousands of galaxies tightly concentrated on the plane of the sky (see Figure 2.1). The first to independently identify and catalog these concentrations of nebulæ were Messier (1784) and Herschel (1785). In the beginning of the 20th century, Slipher and Hubble established the extragalactic nature of nebulæ, now identified as galaxies. Optical discoveries of clusters increased over the following years, culminating with the Abell (1958) catalogs (see Biviano, 2000, for an historical review).

Most of the techniques employed for optical cluster identification are based on Abell's approach of measuring the richness of a system (e.g. Lumsden et al., 1992; Dalton et al.,

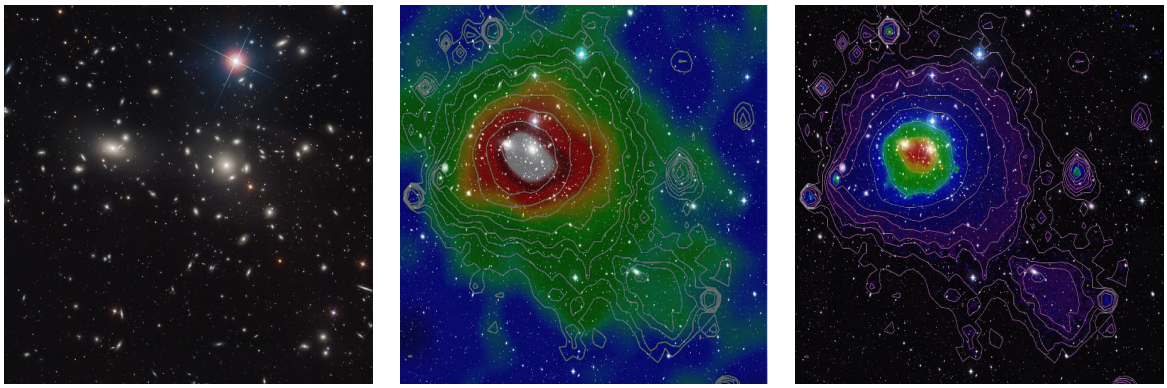

Figure 2.1: The Coma galaxy cluster in optical data from the Sloan Digital Sky Survey (**left** panel), as seen in by Planck through the SZ effect (**central** panel, shaded color), and by ROSAT in X-rays (**right** panel, shaded color and contours in both panels). The images in the central and right panels are overlaid on optical images obtained by DSS. Credits: ESA / LFI and HFI Consortia (Planck image); MPI (ROSAT image); NASA/ESA/DSS2/SDSS (otical images). Acknowledgement: Davide De Martin (ESA/Hubble).

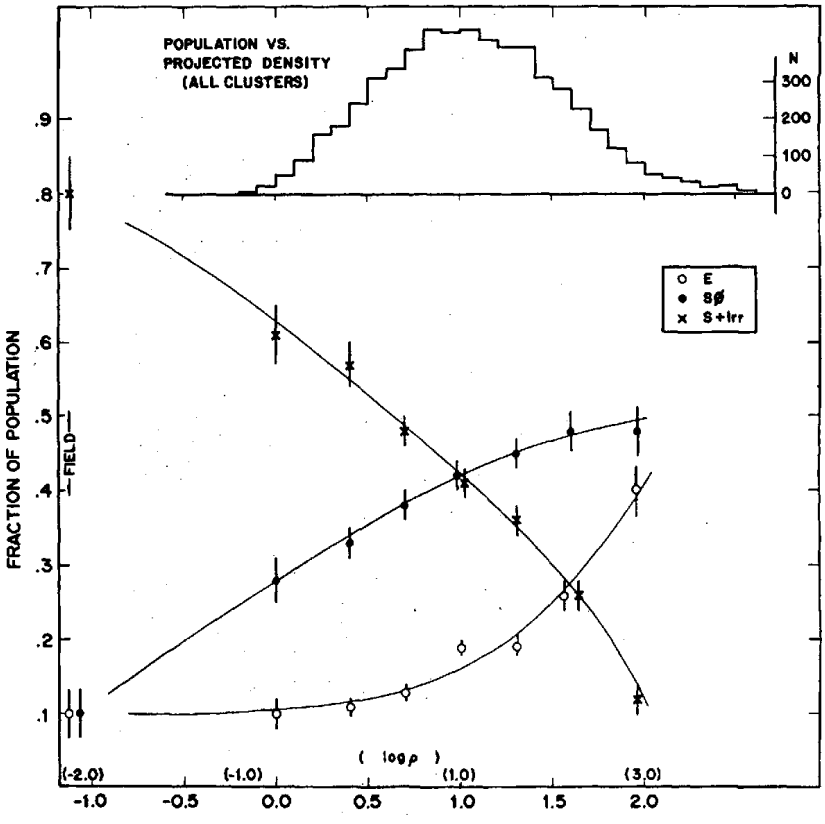

Figure 2.2: The morphology-density relation for 55 galaxy clusters. The fraction of elliptical galaxies (E) increases as a function of local galaxy density, while the spirals one (S+irr) decreases (lower panel). The upper panel is an histogram showing the number distribution. Credits: Dressler (1980).

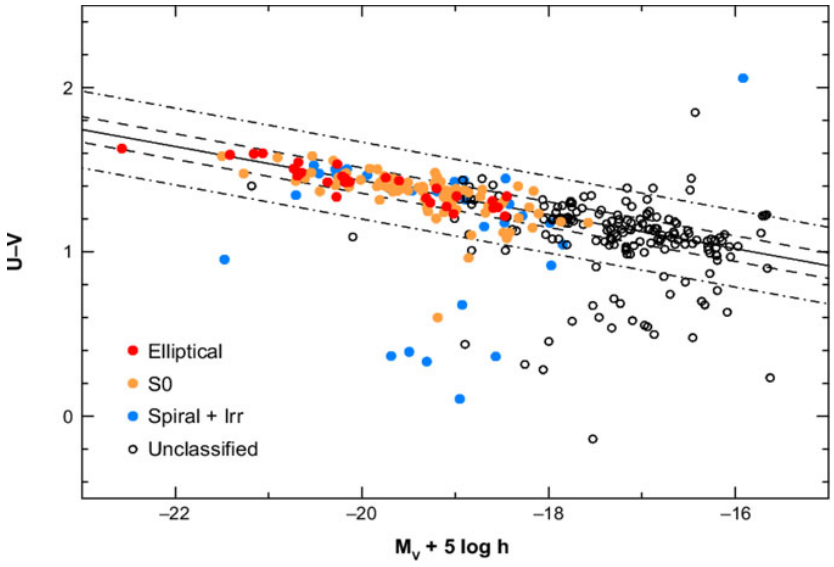

Figure 2.3: The (U-V)-M_V color-magnitude relation for spectroscopic members of the Coma cluster. Symbols represent the different morphological types of galaxies. Credits: Bower et al. (1999).

1997, see Section 2.1.1). These methods were later refined and extended to include information on galaxy colors (Gladders & Yee, 2000; Bahcall et al., 2003; Rykoff et al., 2014). Infrared colors have particularly been useful to allow cluster detection at high redshift ($z \geq 1$) (Papovich, 2008; Eisenhardt et al., 2008; Wilson et al., 2009), where their identification is otherwise challenging.

Studies of galaxy populations have shown a correlation between local galaxy density and morphological and stellar population properties. This connection leads to the well-known morphology-density and star formation rate-density relations. The *morphology-density* relation, illustrated in Figure 2.2, shows that the fraction of elliptical galaxies increases as a function of local galaxy density, while the fraction of spirals decreases. Galaxy clusters are thus mainly populated by elliptical galaxies. Moreover, ellipticals have an old and passive stellar population, characterized by red colors. In a diagram of galaxy color versus magnitude, red member galaxies are placed on a narrow strip known as *red sequence*. Figure 2.3 shows a rendition of the color-magnitude distribution of galaxies in the Coma cluster, with the elliptical galaxies being tightly concentrated on the red sequence.

Mass estimates based on cluster galaxies

Under the assumption that light traces mass, the total luminosity of a cluster is an indicator of its mass. Measuring the luminosity of each cluster galaxy is impractical, and becomes impossible at high redshifts, where only the brightest galaxies can be observed. However, the luminosity distribution of cluster galaxies is nearly the same for all clusters. Thus, observing the high-luminosity tail of that distribution allows one to normalize the luminosity function, yielding estimates for both the cluster's total optical luminosity and its mass. This information is encoded into the definition of *richness*.

Several estimators of richness exist in literature. In his catalogs, Abell measured it as the number of galaxies in the magnitude range m_3 to m_3+2, where m_3 is the magnitude of the third brightest galaxy. Other studies base their definition of richness on the assumption of a defined luminosity distribution function. The luminosity distribution of bright cluster galaxies is well described by the Schechter (1976) function,

$$\frac{\mathrm{d}^2 N}{\mathrm{d}L\,\mathrm{d}V} = n_0 \left(\frac{L}{L^*}\right)^{\alpha} \mathrm{e}^{-L/L^*}, \tag{2.1}$$

with n_0 being a normalization factor. The density of objects below the characteristic luminosity L^* decreases as a power-law with slope α, and is exponentially truncated at high luminosities. The brightest cluster galaxy (BCG) is generally found close to the spatial and kinematic center of the halo.

Postman et al. (1996) define a richness parameter $\Lambda_{\rm cl}$ corresponding to the number of galaxies brighter than the characteristic luminosity L^*. Another richness parameter is $\mathrm{B_{cg}}$, defined as the amplitude of the cluster center–galaxy correlation function, normalized by the Schechter luminosity function (Yee & López-Cruz, 1999). Rykoff et al. (2014) estimate a membership probability ($P_{\rm mem}$) for each galaxy in the vicinity of a cluster, and define the cluster richness as the sum of the membership probabilities over all galaxies $\lambda = P_{\rm mem}$. As detailed in Chapter 4, the richness is a powerful mass-proxy, tightly correlated with the total mass.

The first hint of a relation between richness and velocity dispersions came from a study of Hubble & Humason (1931). They noted that the velocity range spanned by Coma galaxies was larger than in other clusters having fewer galaxy members. Zwicky (1933) saw the great potential of these velocity dispersion measurements, and used them for deriving the mass of the Coma cluster via the virial theorem.

This theorem relates the gravitational potential energy U of a self-gravitating body to the total kinetic energy T of its constituent parts, such that $2T+U=0$. Let us consider a sphere of radius R and mass M, with a small test mass m sitting on its surface. We have

$$U = G\frac{M(R)m}{R}, \tag{2.2}$$

and

$$T = 1/2mv^2 = 1/2m(3\sigma^2_{\rm LOS}), \tag{2.3}$$

where $\sigma_{\rm LOS}$ is the velocity dispersion along the observer's line of sight[1]. Using this method, Zwicky estimated a total mass of the order of $10^{15} M_\odot$. However, measuring the combined luminosity of all the members of the Coma cluster, he found a value of approximately $10^{13} L_\odot$. For an old stellar population with relatively large stellar mass-to-light ratio, this corresponds to a stellar mass of few times $10^{13} M_\odot$. He had discovered the *missing mass* problem. This large discrepancy between the stellar mass and the total mass content has been consistently found in galaxy clusters, and led to the discovery of dark matter (see

[1] For isotropic motions, $v^2 = 3\sigma_{\rm LOS}$.

Section 2.1.3).

Since then, many other techniques have been developed, linking velocity dispersion to the cluster total mass. As an alternative to the virial theorem, Bahcall & Tremaine (1981) introduced the *projected mass estimator*, based on moments of the projected mass. Exploiting the observed distribution of cluster galaxies in projected phase-space, the cluster mass can also be determined via the *caustic* method (Diaferio & Geller, 1997; Diaferio, 1999; Rines & Diaferio, 2006). More recently, the velocity dispersion-mass scaling relation has been calibrated using numerical simulations (e.g. Biviano et al., 2006; Evrard et al., 2008; Saro et al., 2013). Another efficient technique is based on the Jeans equation (e.g. Binney & Tremaine, 1987). This method has been extensively used to recover dynamical masses while also gaining information on galaxy formation and evolution (Biviano & Poggianti, 2009; Biviano et al., 2013, 2017; Munari et al., 2014). The analyses presented in the second part of this thesis are based on this method. We provide a more detailed description of the Jeans analysis in Section 2.3.

Evolution of cluster galaxies

The analysis of the phase-space distributions of cluster galaxies is also a useful tool to constrain models of galaxy formation and evolution. In this context, it is important to distinguish among galaxies of different morphological types, as they may not have formed and evolved in the same way.

In 1940, Holmberg found the first evidence for kinematical segregation of cluster galaxies. He noticed that Virgo spirals had a larger velocity dispersion than ellipticals. Over the years, an increasing number of spectroscopic observations made it possible to confirm this result. In particular, Zwicky & Humason (1964) found also hints of a luminosity segregation, with the fainter, less massive galaxies having a higher velocity dispersion than the brighter ones, indicating a trend towards equipartition of kinetic energy. This result was later confirmed by Rood et al. (1972) and Chincarini & Rood (1977). Tammann (1972) put Holmberg's early result on a solid basis, later extended by Moss & Dickens (1977) to clusters in general. Today, this result has been confirmed by several studies (Biviano et al., 1992; Andreon, 1996; Stein, 1997), which also showed that the velocity dispersion profile of spirals is significantly steeper than that of early-type galaxies (Mohr et al., 1996; Biviano et al., 1997; Carlberg et al., 1997c; Adami et al., 1998).

From the theoretical side, the evidence for segregation is interpreted as arising in a different formation process for galaxies of different morphologies, and/or a different evolutionary history. Many physical processes can affect the morphology, luminosity, and velocity of a galaxy. Dynamical friction (Chandrasekhar, 1943) slows down the more massive galaxies, circularizes their orbits (e.g. den Hartog & Katgert, 1996), and enhances their merger rate (e.g. Mamon, 1992); ram pressure and collisions can transform a star forming galaxy into a quiescent one, also affecting its morphology; tidal effects can truncate the galaxy surface brightness and stellar mass profiles.

On the other hand, the steeper and larger velocity dispersion of star-forming galaxies suggests that they are falling into the cluster (Biviano et al., 1997). The hierarchical model for the formation of cosmological structures predicts the infall of field galaxies into galaxy systems. This behavior has been observed in a study of the pattern of peculiar velocities around nearby groups (Ceccarelli et al., 2005). Further evidence can be obtained from the analysis of the mass profiles out to larger radii. Rines & Diaferio (2006) found that, on average, the mass within the turnaround radius[2] of nearby clusters is about twice the mass within their virial radius. These objects will thus keep accreting material from the surrounding regions, eventually doubling their virial mass. Biviano & Katgert (2004) determined the velocity anisotropy profiles of different classes of galaxy types in nearby clusters, finding that early-type galaxies move on isotropic orbits ($\beta \sim 0$) at all radii, while late-type ones have increasingly radial orbits. This behavior is similar to that obtained for dark matter particles in numerical simulations (e.g. Ghigna et al., 1998; Diaferio, 1999), This is a hint that late-type galaxies, having been recently accreted into the cluster, still retain memory of (mostly radial) gravitational infall along the filaments connecting to the cluster.

Several studies have been focused on exploiting the phase-space distribution of cluster galaxies to gain information on their formation and evolution (Biviano & Poggianti, 2009; Biviano et al., 2013, 2017; Munari et al., 2014). In Chapter 3 we present a study of the orbital properties of 110 galaxy clusters selected using the Sunyaev-Zel'dovich effect (SZE).

2.1.2 Intra-cluster medium (ICM)

Galaxies only make up for $\sim 5\%$ of the clusters' baryonic matter. The remaining $\sim 15\%$ is in the form of hot gas ($\sim 10^7 - 10^8$K), mainly hydrogen and helium, with low density ($\sim 10^{-3}\,\mathrm{atoms/cm^3}$). The intra-cluster gas is observable through two signatures: the thermal bremsstrahlung and the Sunyaev-Zel'dovich (SZ) effect.

Thermal bremsstrahlung

Given its high temperature and low density, the ICM is almost completely ionized. Electrons and nuclei are in a state of collisional equilibrium. Gas collisions produce the emission of bremsstrahlung photons at X-ray energies. The spectral energy distribution of the bremsstrahlung is very sensitive to the gas temperature, thus allowing its determination. The bolometric emissivity of the thermal bremsstrahlung takes the form

$$\epsilon_{\mathrm{bol}} \propto \sqrt{\frac{T}{1\mathrm{KeV}}} \left(\frac{n_{\mathrm{e}}}{10^{-3}\,\mathrm{cm^{-3}}}\right)^2 \mathrm{erg\ s^{-1} cm^{-3}}, \tag{2.4}$$

where n_{e} is the number density of electrons, and we assumed a fully ionized plasma. We note that the emissivity is proportional to the square of gas density, making galaxy clusters clear X-ray sources in the sky and the gas density constrainable.

[2]The turnaround radius marks the region where the Hubble flow balances the infall motion.

Assuming that the ICM is spherically symmetric and in hydrostatic equilibrium in the cluster, the potential and thus the cluster mass can be derived. With these assumptions, the gravitational pull due to the enclosed mass $M(< R)$ inside any radius r will be balanced by the gas pressure gradient,

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{GM(<r)\rho_{\mathrm{gas}}(r)}{r^2}. \tag{2.5}$$

Keeping in mind the ideal gas law, $p = \rho_{\mathrm{gas}} k_{\mathrm{B}} T/\mu m_{\mathrm{p}}$, we can solve for $M(< R)$, finding the enclosed mass to be related to the observables, the gas density ρ_{gas} and temperature T, as (Sarazin, 1988):

$$M(<r) = -\frac{k_{\mathrm{B}} T(r)\, r}{G\mu m_{\mathrm{p}}} \left(\frac{\mathrm{d}\ln \rho_{\mathrm{gas}}}{\mathrm{d}\ln r} + \frac{\mathrm{d}\log T(r)}{\mathrm{d}\ln r} \right), \tag{2.6}$$

where G is the gravitational constant, k_{B} the Boltzmann's constant, m_{p} the proton mass, and μ the mean molecular weight of the gas.

The assumption of hydrostatic equilibrium requires a stationary gravitational potential on a sound crossing time, no relevant contributions from forces other than gas pressure and gravity, and only thermal pressure to support the cluster gas from collapse. The application of this assumption to systems undergoing major merger events, in regions with strong AGN feedback (e.g. Fabian et al., 2003; McNamara & Nulsen, 2007), or with increased levels of non-thermal pressure support (e.g. Nagai et al., 2007; Mahdavi et al., 2008; Shi & Komatsu, 2014) causes an intrinsic dispersion of the gas observables at fixed mass. On the other hand, X-ray masses are relatively insensitive to triaxiality (Gavazzi, 2005).

The hydrostatic method can be implemented using a number of different approaches. The most common one consists in using parametric functions to model the gas density and temperature profiles through the observed surface brightness and spectral temperature data. These are then used to calculate the partial derivatives at each radius, determining the total mass profile via Equation 2.6 (Cavaliere & Fusco-Femiano, 1976; Pratt & Arnaud, 2002; Vikhlinin et al., 2006). This method, however, imposes the use of strong priors. A second approach employs a functional form of the gravitational potential. Together with the geometrical deprojection of the surface brightness, this is then used to recover a temperature profile via Equation 2.6. The modeled T profile is then fitted to the observed one to minimize a merit function that depends only on the parameters describing the mass model (Allen et al., 2008; Rasia et al., 2006; Meneghetti et al., 2010; Ettori et al., 2013; Amodeo et al., 2016).

Over the years, several studies reported mismatches between hydrostatic mass estimates and mass estimates derived by alternative means (e.g. Miralda-Escude & Babul, 1995; Girardi et al., 1997; Ota et al., 2004; Arnaud et al., 2007; Lau et al., 2009). A possible origin for this tension is that there are deviations from hydrostatic equilibrium. These deviations are commonly quantified via the hydrostatic mass bias parameter, b, defined as the fractional difference between the true mass and that inferred by assuming hydrostatic

equilibrium (see, e.g., Donahue et al., 2014; von der Linden et al., 2014b; Hoekstra et al., 2015; Biffi et al., 2016; Hurier & Angulo, 2018).

Sunyaev-Zel'dovich effect

The second observational signature, known as (thermal) Sunyaev-Zel'dovich effect (SZE, Sunyaev & Zeldovich, 1972), is created by CMB photons passing through a cluster. A fraction of these low-energy photons interacts with the high-energy electrons in the hot ICM via inverse Compton scattering, producing a characteristic distortion of the CMB spectrum at the location of the cluster.

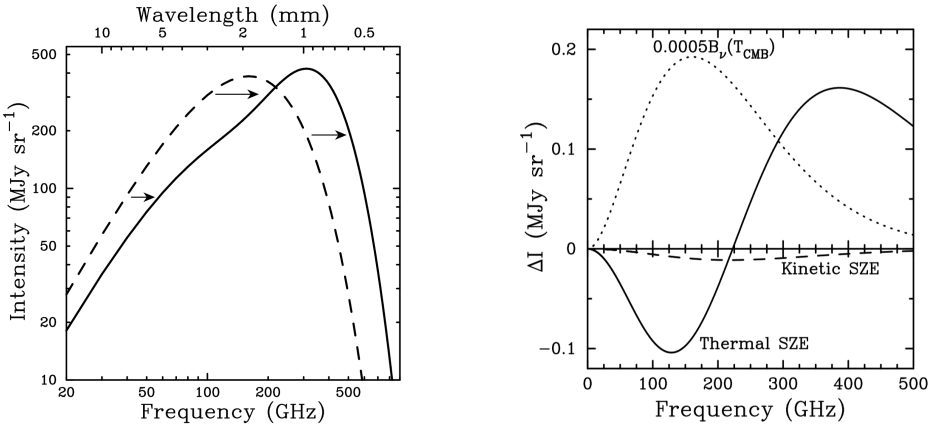

Figure 2.4: The CMB spectrum (dashed), distorted by the interactions with the hot ICM (thermal SZ effect). This effect appears as a decrease in the intensity of the CMB at frequencies $\lesssim 218$ GHz, and as an increase at higher frequencies. For illustration purposes, the amplitude of the effect has been strongly enhanced. Credits: Carlstrom et al. (2002).

This effect, illustrated in Figure 2.4, appears as a decrease in the intensity of the CMB black-body spectrum at frequencies $\lesssim 217$ GHz, and as an increase at higher frequencies. In high-resolution images of the CMB (see Figure 1.4), galaxy clusters stand out as darker (at frequencies $\lesssim 217$ GHz) or brighter (at higher frequencies) spots. At ~ 217 GHz there is no shift in the CMB intensity. The amplitude of the SZ effect is given by the Compton y-parameter

$$y = \frac{\sigma_\mathrm{T}}{m_\mathrm{e}c^2} k_\mathrm{B} \int n_\mathrm{e} T \mathrm{d}V = \frac{\sigma_\mathrm{T}}{m_\mathrm{e}c^2} \int P \mathrm{d}V, \tag{2.7}$$

where σ_T is the Thomson cross-section, m_e the electron mass, k_B the Boltzmann constant, and n_e the electron number density.

From Eq. 2.7 we see that the Compton y-parameter, at fixed thermal energy, does not change with cluster redshift. However, the observed signal is given by the integrated Compton y-parameter over the solid angle of a given cluster, $\mathrm{d}\Omega \propto 1/D_A(z)^2$, where $D_A(z)$ is the angular diameter distance defined in Sect. 1.1. At $z \gtrsim 1$ the angular diameter distance $D_A(z)$ is approximately flat. As the universal matter density increases as $(1+z)^3$, a cluster

of given mass is denser, therefore hotter, at higher redshifts, producing a higher SZ signal. As a result, SZE flux limited surveys (i.e. surveys having effective beam size larger than the size of the cluster) will detect all clusters above a detection mass threshold, with the limiting mass only slightly declining at $z \gtrsim 1$. This feature allows us to detect clusters out to the highest possible redshifts, making SZE-selected cluster samples particularly suitable for studies of cluster formation and growth of structure.

The SZE signature has been observed by surveys carried out by, e.g., the South Pole Telescope (SPT, Carlstrom et al., 2011), the Atacama Cosmology Telescope (ACT, Fowler et al., 2007), and the Planck satellite (Planck Collaboration et al., 2011). The central panel of Figure 2.1 shows a map of the surface density of y for the Coma cluster, obtained from CMB observations of the Planck satellite. In Section 3 we present a study based on SPT-SZ selected clusters, describing this survey in more detail.

2.1.3 Dark matter

As mentioned in Section 2.1.1, Zwicky (1933) observed a discrepancy between the gravitating mass of the Coma cluster, obtained from velocity dispersions of its member galaxies, and the mass estimated from the total stellar luminosity. This discrepancy can not be fixed by the addition of the intra-cluster gas, which is found to be about 10% of the total cluster mass (see, e.g., Allen et al., 2002). The dynamics of galaxy clusters thus requires the presence of large amounts of matter not accounted for by visible baryonic matter: *dark matter*.

Many hypotheses were formulated on the nature of dark matter, such as intergalactic material (Rood, 1965; White, 1977), massive collapsed objects (Van den Bergh, 1969), HI *snowballs* (Peebles, 1971), and M8 dwarf stars (Tarter & Silk, 1974). Most of these possibilities were eventually ruled out by measurements of the abundance of light elements produced in the primordial nucleosynthesis, favoring a non-baryonic, electromagnetically not interacting dark matter. This result has been confirmed with higher precision by measurements of the CMB.

Nowadays, it is still not clear what constitutes dark matter. The most popular candidates are weakly interacting massive (GeV scale) particles (WIMPs). Many experiments are attempting to directly and indirectly detecting such particles (Angloher et al., 2012; Aguilar et al., 2013; Roszkowski et al., 2018). The leading dark matter alternative to the WIMPs is axions: low mass, weakly interacting particles with zero spin, emerging from a theory that was proposed to solve the strong charge-parity (CP) problem (Lyapustin, 2011; Du et al., 2018).

Numerical simulations with collisionless, cold dark-matter particles show that, on average, dark matter halos are well described by an universal density profile, the Navarro, Frenk, & White (1997, hereafter NFW) profile:

$$\rho(r) = \frac{\rho_0}{(r/r_\mathrm{s})(1 + r/r_\mathrm{s})^2}, \tag{2.8}$$

with the two parameters ρ_0 and r_s corresponding to a scale in density and radius. The scale radius relates to the virial radius R_{200} via the concentration $c = R_{200}/r_s$. The concentration is correlated with the halo mass and redshift. Wechsler et al. (2002) show that the concentration strongly correlates with the formation epoch. In fact, the density in clusters cores reflects the background density at formation time. More massive clusters form later, and are therefore expected to have lower concentration. Several studies aimed at investigating the concentration–mass relation, assuming different cosmologies, halo density profiles and resolutions (e.g. Duffy et al., 2008; Macciò et al., 2008a; Meneghetti et al., 2014), often modeling this relation as

$$c(M, z) \propto A\, M^B\, (1+z)^C, \tag{2.9}$$

with fitting parameters A, B and C.

2.2 Gravitational lensing

General relativity predicts that the gravity associated with a distribution of matter will bend light rays passing close to it[3]. This phenomenon is known as *gravitational lensing* (see Bartelmann, 2010, for a recent review). Gravitational lensing happens on all scales – photons emitted by distant galaxies can be bent by galaxies and clusters of galaxies, as well as stars and planets. As if we are viewing these sources through a piece of glass with a spatially varying index of refraction, their images will appear slightly distorted and magnified by this effect.

Figure 2.5 shows the the geometry of a typical gravitational lens system. A light ray is emitted from a source S and, on its way to the observer O, is deflected at the lens plane by an angle $\vec{\hat{\alpha}}$. The angle between the optical axis and the image I is denoted by $\vec{\theta}$, while $\vec{\beta}$ is the angle between the optical axis and the true source. D_s, D_d and D_{ds} are the angular diameter distances between observer and source, observer and lens, lens and source, respectively.

The deflection angle $\vec{\hat{\alpha}}$ is related to the projected gravitational potential of the lens, such that

$$\vec{\hat{\alpha}} = 2 \int \nabla_\perp \Phi \mathrm{d}s, \tag{2.10}$$

where ds is the photon path. This angle is related to the observed deflection angle as:

$$\vec{\alpha} = \frac{D_{ds}}{D_s} \vec{\hat{\alpha}}. \tag{2.11}$$

The observed image and true source positions are related through the *lens equation*:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \tag{2.12}$$

[3]Newtonian gravity also predicts gravitational lensing, but smaller by a factor 2.

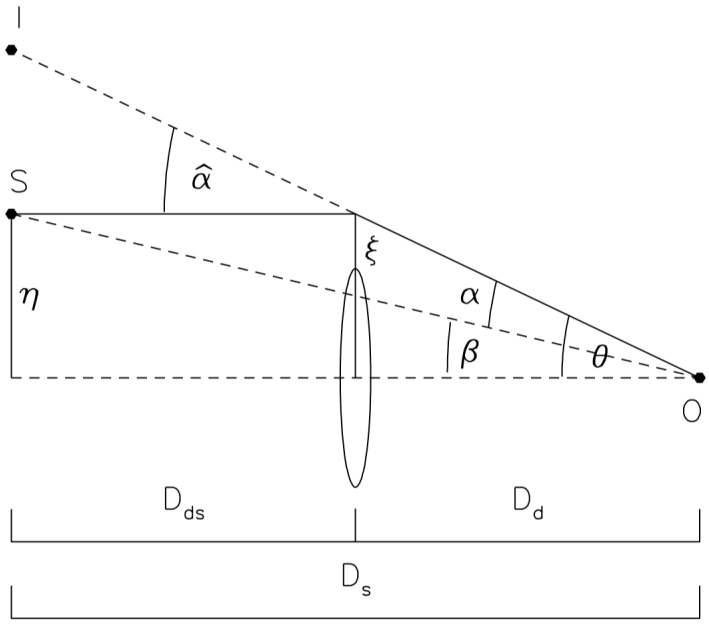

Figure 2.5: Representation of a typical gravitational lensing diagram. D_s, D_d and D_{ds} are the angular diameter distances between observer (O) and source (S), observer and lens, lens and source, respectively. The source of the light ray is at a transverse distance $\vec{\eta}$ from the optical axis to the observer O, crossing the lens plane at transverse distance $\vec{\xi}$, deflected by an angle $\vec{\alpha}$. The angular separations of the observed image (I) and the source from the observer optical axis are $\vec{\theta}$ and $\vec{\beta}$, respectively. Credits: Narayan & Bartelmann (1996).

According to the positions of the source, lens and observer, and the mass and shape of the lens, we can distinguish between two different regimes: strong lensing and weak lensing. If the deflection angle is comparable to the angular size of the source, we have a *strong lensing* effect. This will lead to strong distortions and multiple images of the source, which sometimes blend together to produce arcs (see Figure 2.6). Otherwise, if $\vec{\alpha}$ is much smaller than the source size, only a small deformation is produced (*weak lensing*). In this case, the source will still be distorted – both stretched (shear) and magnified (convergence).

For a galaxy cluster and background galaxies of known redshifts, the measured gravitational shear can be used to infer the cluster mass. In fact, the amplitude of the distortion provides a direct measure of the projected gravitational potential, independent of the nature of the dark matter or the dynamical state of the system. The more massive the object, the stronger its gravitational field and hence the greater the bending of light rays. The most common technique to calibrate weak lensing masses is based on fitting the observed gravitational shear profile with a parametrized mass model (e.g. Hoekstra, 2007). Analyses of galaxy clusters have also been used to calibrate mass–observable scaling relations (McClintock et al., 2019; Dietrich et al., 2019) and to constrain cosmological parameters (Mantz et al., 2014; Bocquet et al., 2018). Strong lensing, on the other hand, enables mass measurements of the regions enclosed by gravitational arcs. These constraints, combined

Figure 2.6: The region of sky around Abell 2218, as seen by the Hubble telescope. The gravitational potential of this massive galaxy cluster deflects the light from background galaxies, producing deflections, distortions and multiple images. Credits: NASA, ESA, and Johan Richard (Caltech, USA).

with those from weak lensing, can significantly improve the calibration of projected cluster masses (e.g. Meneghetti et al., 2010). Strongly-lensed clusters have also been used to constrain the geometry of the Universe (Jullo et al., 2010), as has been done with the shear ratio test for weak lensing measurements (Taylor et al., 2007).

As for the X-ray mass measurements, the lensing method suffers from several assumptions and problems, such as the mass-sheet degeneracy (Dye et al., 2001; Cypriano et al., 2004), projection effects (Metzler et al., 2001; Wambsganss et al., 2005), and low-z inefficiency (Natarajan & Kneib, 1997). On the other hand, no assumptions are made on the dynamical state of the gravitating matter (Bartelmann, 2010).

2.3 The Jeans analysis

The analyses presented in Chapters 3, 4, and 5 are based on the use of the Jeans equation as a tool to recover the total cluster mass. This equation is based on the view of galaxies as a system of particles in the six dimensional *phase-space*. In this context, a galaxy is instantaneously described by the *distribution function* $f(\vec{x}, \vec{v})$ over the phase-space, with $\vec{v}$ and $\vec{x}$ being the position and velocity of the galaxy. For the derivations obtained in this section, we follow Binney & Tremaine (1987).

The time evolution of the distribution function is governed by a 6-dimensional continuity equation, analogous to the 3-dimensional equation of continuity of fluid mechanics. Each point in phase-space is described by a 6-D vector $\vec{w} = (\vec{x}, \vec{v})$. The continuity equation is then:

$$\frac{\partial f}{\partial t} + \sum_{\alpha=1}^{6} \frac{\partial}{\partial w_\alpha}(f \dot{w_\alpha}) = 0. \tag{2.13}$$

We can see that

$$\sum_{\alpha=1}^{6} \frac{\partial \dot{w_\alpha}}{\partial w_\alpha} = \sum_{i=1}^{3} \left(\frac{\partial v_i}{\partial x_i} + \frac{\partial \dot{v_i}}{\partial x_i} \right) = 0. \tag{2.14}$$

The first part of this sum is zero, because the velocities are not explicit functions of position. From the equation of motion we have that $\dot{v_i} = -\frac{\partial}{\partial x_i}\Phi(\vec{x})$. The continuity equation thus becomes

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f - \vec{\nabla}\Phi \cdot \frac{\partial f}{\partial \vec{v}} = 0. \tag{2.15}$$

which is the *collisionless Boltzmann equation* (CBE)[4].

Finding a distribution function compatible with a given density distribution is not straightforward, and often does not yield an unique solution. On the other hand, it is relatively easy to calculate the moments of f. We can thus collapse the 6-dimensional phase-space density into a set of functions of 3-dimensional positions by taking moments of the velocities. The zeroth moment is the number density,

$$\nu(\vec{x}) \equiv \int f(\vec{x}, \vec{v}) \mathrm{d}^3\vec{v}. \tag{2.16}$$

For each of three velocity components, the first moment gives a mean velocity,

$$\vec{v_i}(\vec{x}) \equiv \frac{1}{\nu(\vec{x})} \int v_i f(\vec{x}, \vec{v}) \mathrm{d}^3\vec{v}. \tag{2.17}$$

One can likewise define higher order moments with combinations of powers of the three velocity components. The second moments give a quantity related to the velocity dispersion

[4]It is collisionless in the sense that particles do not make instantaneous jumps in $\vec{x}$ or $\vec{v}$, a consequence of a potential Φ that is smooth in space and time.

tensor, σ_{ij}^2,

$$\overline{v_i v_j}(\vec{x}) \equiv \frac{1}{\nu(\vec{x})} \int v_i v_j f(\vec{x}, \vec{v}) \mathrm{d}^3\vec{v} = \sigma_{ij}^2 + \overline{v}_i \overline{v}_j, \tag{2.18}$$

with $i, j = 1, 2, 3$.

By multiplying the CBE by powers of the velocity components, and integrating over velocity space, we can obtain the differential equations for the various velocity moments. The zeroth moment of the CBE yields

$$\int \frac{\partial f}{\partial t} \mathrm{d}^3\vec{v} + \int v_i \frac{\partial f}{\partial x_i} \mathrm{d}^3\vec{v} - \frac{\partial \Phi}{\partial x_i} \int \frac{\partial f}{\partial v_i} \mathrm{d}^3\vec{v} = 0, \tag{2.19}$$

where we dropped the summation signs and adopted the implicit summation over i. We can eliminate the last term by applying the divergence theorem,

$$\int \vec{g} \cdot \vec{\nabla}_v f \mathrm{d}^3\vec{v} = \oint f \vec{g} \cdot \mathrm{d}\vec{S} - \int f \vec{\nabla}_v \cdot \vec{g} \, \mathrm{d}^3\vec{v}. \tag{2.20}$$

In the present case, $g = \vec{\nabla}\Phi(\vec{x})$. As g is not a function of $\vec{v}$, we can move it outside the integral. Moreover, $\vec{\nabla}_v \cdot \vec{g} = 0$, and the surface integral goes to zero for a phase-space density that goes to zero at infinity. With this in mind, and incorporating Equations 2.16 and 2.17, we have

$$\frac{\partial}{\partial t}\nu + \frac{\partial}{\partial x_i}(\nu \overline{v}_i) = 0. \tag{2.21}$$

which has the form of the standard 3-D continuity equation.

The first moment of the CBE is found by multiplying Eq. 2.15 by v_j and integrating over the velocity:

$$\frac{\partial}{\partial t} \int v_j f \mathrm{d}^3\vec{v} + \frac{\partial}{\partial x_i} \int v_j v_i f \mathrm{d}^3\vec{v} - \frac{\partial \Phi}{\partial x_i} \int v_j \frac{\partial f}{\partial v_i} \mathrm{d}^3\vec{v} = 0. \tag{2.22}$$

Integrating by parts and expressing our results in terms of average velocities, we get

$$\frac{\partial}{\partial t}(\nu \overline{v}_j) + \frac{\partial}{\partial x_i}(\nu \overline{v_j v_i}) + \frac{\partial \Phi}{\partial x_i} \int f \frac{\partial v_j}{\partial v_i} \mathrm{d}^3\vec{v} = 0. \tag{2.23}$$

In an orthogonal coordinate system, $\dfrac{\partial v_j}{\partial v_i} = \delta_{ij}$, so the last term on the left hand side becomes $\nu\delta_{ij}$. Applying the product rule and the continuity equation to the first term, we get

$$\nu \frac{\partial \overline{v}_j}{\partial t} - (\overline{v}_j) \frac{\partial}{\partial x_i}(\nu \overline{v}_j) + \frac{\partial}{\partial x_i}[\nu(\sigma_{ij}^2 + \overline{v}_i \overline{v}_j)] = -\nu \overline{v}_i \frac{\partial \Phi}{\partial x_j}, \tag{2.24}$$

where we have made use of the relation between the second moments and the velocity dispersion. This leads to an analog of Euler's equation of fluid flow, the *Jeans equations*

for a collisionless fluid:

$$\underbrace{\nu\frac{\partial\overline{v}_j}{\partial t}}_{\text{acceleration}} + \underbrace{\overline{v}_i\nu\frac{\partial\overline{v}_j}{\partial x_i}}_{\substack{\text{kinematic}\\ \text{viscosity}}} = \underbrace{-\nu\frac{\partial\Phi}{\partial x_j}}_{\text{gravity}} - \underbrace{\frac{\partial}{\partial x_i}(\nu\sigma_{ij}^2)}_{\text{pressure}}. \tag{2.25}$$

By setting the acceleration and viscosity terms to zero, we can recover the equation for hydrostatic equilibrium.

It is convenient to express the Jeans equations in spherical coordinates:

$$\frac{\mathrm{d}(\nu\sigma_r^2)}{\mathrm{d}r} + \frac{\nu}{r}\left[2\sigma_r^2 - (\sigma_\theta^2 + \sigma_\phi^2)\right] = -\nu\frac{\mathrm{d}\Phi}{\mathrm{d}r}, \tag{2.26}$$

where $\nu(r)$ is the number density profile of the tracer galaxy population, and $\sigma_r = \overline{v_r^2}$, $\sigma_\theta = \overline{v_\theta^2}$, $\sigma_\phi = \overline{v_\phi^2}$ are the components of the velocity dispersion along the three spherical coordinates r, θ, ϕ. We now introduce the *velocity anisotropy parameter*

$$\beta \equiv 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2}. \tag{2.27}$$

This parameter quantifies the system's degree of radial anisotropy: if all orbits are circular, $\sigma_r = 0$ and $\beta \to -\infty$; if the orbits are isotropic, $\beta = 0$; if they are perfectly radial, $\sigma_\theta = \sigma_\phi = 0$ and $\beta = 1$. For a spherically symmetric potential Φ, we have $\sigma_\theta = \sigma_\phi$, and we obtain

$$\frac{\mathrm{d}(\nu\overline{v_r^2})}{\mathrm{d}r} + 2\frac{\beta}{r}\nu\overline{v_r^2} = \nu\frac{\mathrm{d}\Phi}{\mathrm{d}r}. \tag{2.28}$$

The mass enclosed within a radius r is thus given by

$$M(r) = -\frac{r\sigma_r^2}{G}\left(\frac{\mathrm{d}\ln\nu}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln\sigma_r^2}{\mathrm{d}\ln r} + 2\beta\right). \tag{2.29}$$

Assuming parametric models for the mass and anisotropy profile, fitting the observed distribution of galaxies as a function of clustercentric radii and velocities, it is thus possible to recover the total cluster mass. As detailed in Chapter 3, one of the advantages of calibrating cluster masses through the Jeans analysis is the fact that it allows us to learn about the orbital distribution of the galaxies within clusters (van der Marel et al., 2000a; Biviano & Katgert, 2004), an important piece of information for constraining models of galaxy evolution in clusters. We also show that this method provides tight constraints on the mass estimates ($\sim 15\%$ uncertainty, for cluster stack samples containing ≈ 600 member galaxies).

As for the previously discussed methods to estimate cluster masses (X-ray data, SZE observations, lensing effects), also the Jeans analysis presents some problems and complications. Being based on the assumption of dynamical equilibrium, it is only applicable to virialized systems. As clusters grow by accretion of field galaxies (Moss & Dickens, 1977; Biviano et al., 1997), they are not in a steady state. However, the infall rate of field galaxies into clusters decreases over time, and is estimated to be small for nearby clusters (Ellingson et al., 2001). Moreover, clusters undergoing substantial mass accretion can be identified through the presence of substructures in their phase-space distribution, and excluded from the sample (van der Marel et al., 2000a; Biviano & Girardi, 2003; Katgert et al., 2004a). Clusters undergoing major mergers can also be identified in X-ray observations (Mohr et al., 1993; Buote & Tsai, 1995; Nurgaliev et al., 2013).

Another complication is due to the presence of interlopers, i.e. background/foreground galaxies having velocities and positions in the range of the phase-space distribution of cluster members. Over the years, increasingly sophisticated methods have been developed to identify interlopers (Yahil & Vidal, 1977; den Hartog & Katgert, 1996; Fadda et al., 1996; Carlberg et al., 1997a; Katgert et al., 2004a; Łokas et al., 2006). In the analyses presented in the second part of this thesis, we implement the *Clean* method (Mamon et al., 2013), which identifies these interlopers on the basis of their phase-space location. However, interlopers outside the virial radius, characterized by small peculiar velocities, will be projected inside the projected distribution of member galaxies and will not be identified with this method. Saro et al. (2013) performed an analysis on cosmological N-body simulations showing that, for a population of passive galaxies in massive clusters ($M_{200c} \geq 10^{14} M_{\odot}$) within $3\,R_{200}$, there is a contamination of $\sim$20%. Mamon et al. (2010a) analyzed hydrodynamical simulations, where they found that, after applying an iterative $2.7\sigma_{\rm LOS}$ velocity cut on the projected phase-space distribution of galaxies, the fraction of interlopers is $\sim$ 23% of all DM particles with projected radii within the virial radius. Further studies are needed to better characterize the effects of this systematic on the results of a dynamical analysis.

2.4 Clusters as cosmological probes

Since Zwicky's discovery of dark matter in the Coma cluster (Zwicky, 1933), galaxy clusters have been used as laboratories for cosmological studies. BCGs have been employed as standard candles to investigate the local expansion history of the Universe (Hoessel et al., 1980). In the 1990s, measurements of the baryon fraction in rich clusters were found to be larger than what predicted by cosmic nucleosynthesis for an Einstein-de Sitter Universe (Fabian, 1991; Briel et al., 1992). This contradiction was resolved with the introduction of non-zero cosmological constant (White et al., 1993c). More evidence for the need of dark energy in cosmology was obtained by the discovery of massive clusters at high redshifts (Donahue et al., 1998; Bahcall & Fan, 1998), preceding the measurement of the accelerated expansion of the Universe by Type Ia supernova surveys. In the 2000s, cluster counts started to be used as proxies to measure the amplitude of the matter power spectrum (Borgani et al., 2001; Reiprich & Böhringer, 2002; Allen et al., 2003; Henry et al., 2009).

Thus, clusters provide the means to constrain the parameters of a cosmological model, shed light on the nature of dark matter and dark energy, and test for deviations from GR (Voit, 2005; Allen et al., 2011). In this Section we investigate the number density of clusters as a function of mass, namely the *halo mass function*, and its cosmological implications.

The halo mass function

The halo mass function (MF) is defined as the number density of virialized structures of given mass and redshift,

$$\mathrm{d}N = n(M)\mathrm{d}M. \tag{2.30}$$

In 1974, Press & Schechter proposed an analytical model to calculate $n(M)$. Let's consider a (linearly evolved) density fluctuation field $\delta(\vec{x}, R) \equiv \delta_M$ at the present time, smoothed by a spherical top-hat filter of comoving radius R corresponding to a mass $M = \frac{4}{3}\pi\rho_\mathrm{c}\Omega_\mathrm{m}R^3$, such that

$$\delta_{W_R} = \int \delta_\mathrm{lin}(\vec{x'})W(|\vec{x'} - \vec{x}|; R)\mathrm{d}^3x', \tag{2.31}$$

with

$$W_R(r) = \begin{cases} \dfrac{3}{4\pi R^3}, & r \leq R \\ 0, & r > R \end{cases},$$

Assuming that the primordial density field was Gaussian, the density contrast δ_M inherits this property as well. Thus, the probability distribution of fluctuations between δ_M and $\delta_M + \mathrm{d}\delta_M$ is given by

$$P(\delta_M)\mathrm{d}\delta_M = \frac{1}{\sqrt{2\pi\sigma_M^2}} \exp\left(-\frac{\delta_M^2}{2\sigma_M^2}\right)\mathrm{d}\delta_M. \tag{2.32}$$

Consequently, the probability that the linearly evolved δ_M exceeds the critical density contrast $\delta_\mathrm{vir,lin}$ (Eq. 1.83) is given by

$$P_{\delta_M > \delta_\mathrm{vir,lin}} = \int_{\delta_\mathrm{vir,lin}}^{\infty} P(\delta_M)\mathrm{d}\delta_M = \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_\mathrm{vir,lin}}{\sqrt{2}\sigma_\mathrm{M}}\right), \tag{2.33}$$

where erfc is the complementary error function, and σ_M is given by Eq. 1.70, inserting the Fourier transform of the filter function $W_R(r)$.

Equation 2.33 thus represents the fraction of matter virialized in halos of mass M or larger. As we go to smaller masses, we would expect this to give the total fraction of mass in the Universe that is locked up in bound objects. However, from Eq. 1.71 with $n > -3$ we have that $\sigma_M \to \infty$ for $M \to 0$. Hence, the argument of erfc goes to zero, erfc(0)=1, and the probability turns out to be 1/2. Only 50% of the mass of the Universe resides in bound structures. This is known as the *cloud-in-cloud problem*, arising from incorrectly accounting for underdense regions. These underdense regions, corresponding to half the mass, could, at a later time, be contained within another object, on a larger mass scale.

Press and Schechter solved this problem by simply introducing a *fudge factor* 2, with the vague understanding that this represents accretion from the underdense regions onto the dense ones.

The halo mass function can thus be written as

$$n(M,z)\mathrm{d}M = \frac{\rho_m}{M}\frac{2\partial P_{\delta_M>\delta_{\mathrm{vir,lin}}}}{\partial M}\mathrm{d}M, \tag{2.34}$$

where the first term, ρ_m/M, is the number density of objects with mass M, while the second term gives the fraction of these objects in the mass interval. Explicitly, we have

$$\begin{aligned} n(M,t)\mathrm{d}M &= 2\frac{\rho_m}{M}\frac{\partial}{\partial M}\left[\frac{1}{2}\mathrm{erfc}\left(\frac{\delta_{\mathrm{vir,lin}}}{\sqrt{2}\sigma_{\mathrm{M}}}\right)\right]\mathrm{d}M \\ &= \sqrt{\frac{2}{\pi}}\frac{\rho_m}{M^2}\frac{\delta_{\mathrm{vir,lin}}}{\sigma_M}\exp\left(-\frac{\delta^2_{\mathrm{vir,lin}}}{2\sigma^2_M}\right)\left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right|\mathrm{d}M. \end{aligned} \tag{2.35}$$

The mass function is very sensitive to cosmological parameters via the variance σ_M, which in turn depends on the cosmological density parameters and the power spectrum. However, the Press and Schechter mass function, while reasonably describing the observed abundance of halos, fails to predict the abundances of halos with high precision. Namely, this framework underestimates the number of low mass halos and over-predicts the abundances of high mass halos. Refined versions of mass functions have been developed thereafter (e.g. Sheth & Tormen, 1999).

The most common way of constructing accurate mass functions at present is based on identifying and counting halos in numerical cosmological simulations and deriving a fitting formula for $n(M,z)\mathrm{d}M$ as a function of mass and redshift. One common functional form is given by Tinker et al. (2008)

$$\frac{\mathrm{d}n}{\mathrm{d}M} = f(\sigma_M)\frac{\rho_m}{M}\frac{\mathrm{d}\ln\sigma_M^{-1}}{\mathrm{d}M}, \tag{2.36}$$

where

$$f(\sigma_M) = A\left[\left(\frac{\sigma_M}{b}\right)^{-a}+1\right]\exp\left(-\frac{c}{\sigma^2_M}\right) \tag{2.37}$$

depends strongly on the peak height. These parameters have to be calibrated on numerical simulations.

Figure 2.7 illustrates the Press and Schechter (dashed) and Tinker (solid line) halo mass functions, at $z = 0$, for three different cosmologies. Differences become most apparent at the high mass end, where the mass functions drop exponentially. Higher values of matter density correspond to a significantly higher density of halos at all mass scales.

Current cosmological constraints

In this section we present some of the latest results obtained from cluster cosmology.

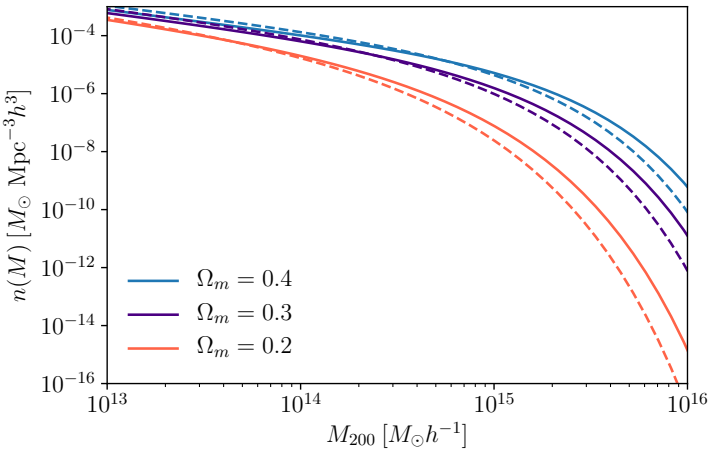

Figure 2.7: The Press and Schechter (dashed) and Tinker (solid lines) halo mass functions, at $z = 0$, for different cosmologies. Note how differences become most apparent at the high mass end, where the mass functions drop exponentially.

The left panel of Figure 2.8 shows constraints on the dark energy equation of state parameter w, from cluster counts and other cosmological probes. As mentioned earlier, cluster abundance is affected by the geometry of the Universe, as well as by the growth of structure. Since both are affected by a change in w, clusters provide some of the tightest single-probe constraints on dark energy. In addition, clusters are complementary to other cosmological probes so a combination yields stronger constraints (Mantz et al., 2015). The right panel illustrates constraints on σ_8 and Ω_m for different versions of the scaling relations between Compton distortion parameter and cluster mass. Blue contours are constraints from CMB anisotropies. Gray contours show results obtained from Weighing the Giants (WtG, von der Linden et al., 2014a), based on gravitational shear information for clusters from Planck Collaboration et al. (2014a). Green contours show results from gravitational lensing of the Canadian Cluster Comparison Project (CCCP, Hoekstra et al., 2015). Violet contours refer to a mass calibration carried out by using lensing of the CMB itself by the clusters (Melin & Bartlett, 2015), and red contours show a recent reanalysis of the CMB lensing data by Zubeldia & Challinor (in prep.). We note how the different definitions of cluster mass employed here affects the results on the cosmological parameters, highlighting the crucial importance of a precise mass calibration.

A range of cosmological constraints coming from observations of clusters, SN Ia (Perlmutter et al., 1999), LSS (Eisenstein et al., 2005), and CMB (Komatsu et al., 2011), find a consistent picture of the Universe dominated by dark energy ($\sim 70\%$), with a sub-dominant dark matter component ($\sim 25\%$), and a small fraction of baryonic material ($\sim 5\%$). Table 2.1 presents a list of the main cosmological parameters of the concordance ΛCDM model, together with their current constraints from CMB measurements as a reference (Planck Collaboration et al., 2016c).

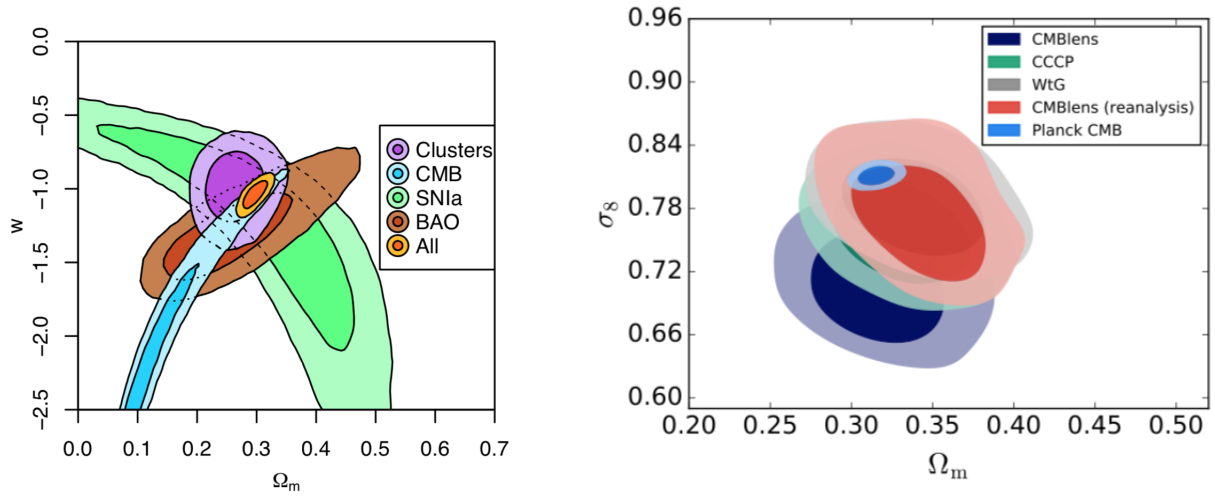

Figure 2.8: **Left:** Constraints in the (Ω_m, w)-plane, for a flat wCDM model. Dark and light shading respectively indicate the 68.3 and 95.4 per cent confidence regions, accounting for systematic uncertainties. Clusters provide the tightest single-probe constraints, made even tighter when combined with other probes. Credits: Mantz et al. (2015). **Right:** Comparison of cosmological constraints in the (Ω_m, σ_8)-plane, at 1 and 2σ, for different versions of the scaling relations between the SZ Compton distortion parameter and cluster mass. Credits: Planck Collaboration et al. (2018a).

Sources of systematic error in cluster cosmology

We now live in an era of precision cosmology, mainly limited by systematic uncertainties (Lima & Hu, 2005; Mantz et al., 2010). Using clusters to determine cosmological parameters requires both an accurate prediction of the halo mass function and a precise measurement of the cluster mass (e.g. from a mass-observable relation, see Section 3.4.4 and Chapters 4 and 5).

A good mass proxy should be straightforward to measure and tightly correlate with mass, exhibiting minimal dispersion across mass and redshift. X-ray observables, such as the luminosity and temperature, present an intrinsic dispersion at fixed mass and redshift of

Symbol	Definition	Constraint
$\Omega_b h^2$	Baryon density	0.02225±0.00016
$\Omega_c h^2$	Cold Dark Matter density	0.1198±0.0015
τ	Reionization optical depth	0.079±0.017
n_s	Primordial scalar spectral index	0.9645±0.0049
H_0	Hubble constant	67.27±0.66
σ_8	RMS matter fluctuations	0.831±0.013

Table 2.1: Parameters of the base ΛCDM cosmology computed from the 2015 baseline Planck likelihoods from CMB measurements (Planck Collaboration et al., 2016c).

$\sim 20-30\%$ (Bulbul et al., 2018; Barnes et al., 2017; Pratt et al., 2009; Lovisari et al., 2015) and $\sim 6-18\%$ (Bulbul et al., 2018; Barnes et al., 2017; Mantz et al., 2010), respectively (see Chapter 5). Optical richness measurements (i.e. the number of galaxies within the detection aperture) also trace the cluster mass. At fixed mass, the scatter in richness is estimated to be $\sim 20\%$ (Rykoff et al., 2012; Saro et al., 2015, see Chapter 4). The Sunyaev-Zel'dovich effect is also an excellent proxy for total cluster mass. The projected, integrated SZ flux or detection signal to noise exhibits an intrinsic scatter of $\sim 20\%$ (e.g., de Haan et al., 2016; Bocquet et al., 2018). Several independent observables can be combined to deliver precise and accurate mass information, and tighter constraints on the cosmological parameters (see, e.g., Bocquet et al., 2015).

This picture is complicated by the presence of measurement related uncertainties and two additional sources of bias, known as *Malmquist* (Malmquist, 1920) and *Eddington* (Eddington, 1913) biases. The first is a selection effect due to the observable threshold of a survey, which hides low-mass objects. The result is an up-scattered observed scaling relation at the low mass end of the observed objects. The Eddington bias, due to the fact that lower mass clusters are more numerous, enhances this effect. Both these systematics can be taken into account by modeling the selection function and the underlying cluster mass function.

Outline and Motivation

In this introduction, we gave an overview of the current state of cluster cosmology, and highlighted the importance of cluster mass calibration for both cosmological and astrophysical studies. We also described a few methods used to estimate the masses of galaxy clusters, with a particular emphasis on the dynamical Jeans analysis. With this background knowledge, we are now ready to discuss the scientific analyses presented in the second part of this thesis.

We begin by presenting a study of galaxy clusters selected using the Sunyaev-Zel'dovich effect (SZE) in the 2500 deg^2 SPT-SZ survey (Chapter 3). As mentioned above, systematic uncertainty in cluster masses is the limiting factor in modern cosmological analyses. In this analysis, we construct a large sample of ∼3000 passive cluster members, spanning the wide redshift range of $0.2 < z < 1.3$. Exploiting this dataset, we perform a dynamical analysis applying the spherically-symmetric Jeans equation, using the cluster member galaxies as tracers of the cluster gravitational potential. The phase-space of cluster galaxies, while providing estimates of the cluster dynamical mass, preserves information about the formation history of the cluster and the evolution of its member galaxies. In this analysis, we exploit all moments of the phase-space distribution (see Section 2.3). First, we analyze the orbital characteristics of the passive galaxies, and their trends with mass and redshift. Then, we investigate the pseudo-phase-space density (PPSD) profile of our clusters to test for dynamical equilibrium and self-similarity. Finally, we perform comparisons with masses obtained through independent calibration in previous analyses. We also investigate the impact of disturbed clusters on our analysis.

In Chapter 4 we present an analysis on the ROSAT All-Sky Survey (RASS) X-ray cluster candidates, which have optical counterparts in SDSS imaging data identified using the redMaPPer algorithm. A subset of these clusters have since been spectroscopically studied within the SPectroscopic IDentification of eRosita Sources (SPIDERS) survey. Our dataset includes 428 CODEX clusters with a corresponding sample of ∼7800 red member galaxies with measured redshifts. For each cluster, we study the λ-$M_{200c} - z$ relation by extracting the likelihood of consistency between the observed phase-space distribution and the modeled projected distribution for a cluster at that redshift and λ.

Finally, in Chapter 5 we extend the calibration of the mass–observable relation to the

X-ray luminosity. In contrast to our study of the richness–mass scaling relation, for this analysis we have to take into account the fact that X-ray sources are likely to be contaminated by random superpositions along the line of sight. We assign each source a probability of being a random superposition, and produce a sample with a 5% contamination fraction, corresponding to roughly a factor of five reduction in contamination. In both our mass–observable scaling relation analyses, we carefully account for the effects of the cluster selection function, statistical corrections for the Eddington and Malmquist biases (see Section 2.4), and additional systematic effects associated with the dynamical mass estimates themselves.

We present our general conclusions in Chapter 6, highlighting the importance of these studies and providing an overview over future and ongoing research projects.

Part II

Original Work

Chapter 3

Galaxy Kinematics and Mass Calibration in Massive SZE Selected Galaxy Clusters to z=1.3

R. Capasso[1,2], A. Saro[1,2,3], J. J. Mohr[1,2,4], A. Biviano[3], S. Bocquet[5,6], V. Strazzullo[1], SPT Collaboration

2019, Monthly Notices of the Royal Astronomical Society (MNRAS), 482, 1043

[1]Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
[2]Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
[4]Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 85748 Garching, Germany
[3]INAF-Osservatorio Astronomico di Trieste via G.B. Tiepolo 11, 34143 Trieste, Italy
[5]Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637
[6]Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL, USA 60439

ABSTRACT

The galaxy phase-space distribution in galaxy clusters provides insights into the formation and evolution of cluster galaxies, and it can also be used to measure cluster mass profiles. We present a dynamical study based on ∼3000 passive, non-emission line cluster galaxies drawn from 110 galaxy clusters. The galaxy clusters were selected using the Sunyaev-Zel'dovich effect (SZE) in the 2500 deg^2 SPT-SZ survey and cover the redshift range $0.2 < z < 1.3$. We model the clusters using the Jeans equation, while adopting NFW mass profiles and a broad range of velocity dispersion anisotropy profiles. The data prefer velocity dispersion anisotropy profiles that are approximately isotropic near the center and increasingly radial toward the cluster virial radius, and this is true for all redshifts and masses we study. The pseudo-phase-space density profile of the passive galaxies is consistent with expectations for dark matter particles and subhalos from cosmological N-body simulations. The dynamical mass constraints are in good agreement with external mass estimates of the SPT cluster sample from either weak lensing, velocity dispersions, or X-ray Y_X measurements. However, the dynamical masses are lower (at the 2.2σ level) when compared to the mass calibration favored when fitting the SPT cluster data to a ΛCDM model with external cosmological priors, including CMB anisotropy data from Planck. The discrepancy grows with redshift, where in the highest redshift bin the ratio of dynamical to SPT+Planck masses is $\eta = 0.63^{+0.13}_{-0.08} \pm 0.06$ (statistical and systematic), corresponding to a 2.6σ discrepancy.

3.1 Introduction

In the current paradigm of structure formation, halos form through the gravitational collapse of overdense regions that are seeded by processes in the early universe. The formation of cold dark matter (CDM) dominated halos proceeds through a sequence of mergers and the accretion of surrounding material, leading to the formation of the galaxy groups and clusters we observe. Baryonic processes associated with the intracluster medium (ICM) and the galaxies also play a role, making galaxy clusters important laboratories for investigations of structure formation and galaxy evolution as well as useful cosmological probes.

Studies of structure formation using cosmological N-body simulations have been used to demonstrate that halos formed from collisionless CDM have, on average, a universal mass density profile (Navarro et al., 1996, 1997, hereinafter NFW). This profile is characterized by two parameters: the virial radius r_{200}[1], and the scale radius r_s, which is the radius at which the logarithmic slope of the density profile is -2. Numerous observational studies have found the mass distributions of clusters to be well described by this model (e.g., Carlberg et al., 1997b; van der Marel et al., 2000a; Biviano & Girardi, 2003; Katgert et al., 2004a; Umetsu et al., 2014).

Another interesting feature is the finding in N-body simulations that the quantity ρ/σ^3, where ρ is the mass density and σ the velocity dispersion, has a power-law form. This quantity is known as pseudo-phase-space density (PPSD) profile, $Q(r)$, and its power-law form resembles that of the self-similar solution for halo collapse by Bertschinger (1985) and is thought of as a dynamical equivalent of the NFW mass density profile (Taylor &

[1] r_Δ defines the sphere within which the cluster overdensity with respect to the critical density at the cluster redshift is Δ. Throughout this paper, we consider $\Delta = 200$ and refer to r_{200} simply as the virial radius.

Navarro, 2001). Others (Austin et al., 2005; Barnes et al., 2006) have suggested that the PPSD profile results from dynamical collapse processes, and should therefore be a robust feature of approximately virialized halos that have undergone violent relaxation (Lynden-Bell, 1967).

The galaxy population is more difficult to study in simulations, because of the over-merging problem, i.e. the premature destruction of dark matter halos in the dense clusters environments in dissipationless N-body simulations (e.g. Moore et al., 1996), and the additional baryonic physics that must be included. However, from the observational side the properties of the galaxy population and trends with mass and redshift can be readily measured and interpreted as long as: (1) selection effects are understood and (2) precise cluster mass measurements are available to ensure that the same portion of the virial region is being studied in clusters of all masses and redshifts. By comparing the galaxy properties to the expectations for collisionless particles studied through N-body simulations, one can characterize the impact of possible additional interactions beyond gravity that are playing a role in the formation of the galaxy population.

As an example, the radial distribution of galaxies in clusters is well fit by an NFW model when clusters are stacked in the space of r/r_{Δ} (e.g. Lin et al., 2004; Muzzin et al., 2007; van der Burg et al., 2014, 2015a; Zenteno et al., 2016). In cluster samples extending to redshift $z \sim 1$, it is clear that the concentration c_{Δ}, defined as the ratio between r_{Δ} and r_s, varies dramatically from cluster to cluster, and that when stacked, the c_{Δ} varies systematically with the prevalence of star formation (Hennig et al., 2017, hereafter H17). The red, passively evolving galaxies have concentrations similar to those expected for the dark matter on these halo mass scales, while the star forming, and presumably infalling blue galaxies are far less concentrated. The number of luminous cluster galaxies (magnitudes $m < m_* + 2$) within the virial region scales with cluster mass as $N \propto M^{\alpha}$ where $\alpha \sim 0.85$ (Lin et al., 2004), and this property appears to be unchanged since redshift $z \sim 1$ (H17). The departure from $\alpha = 1$ in this relation is puzzling, given that massive clusters accrete lower mass clusters and groups (Lin & Mohr, 2004) and this is presumably evidence for galaxy destruction processes that are more efficient in the most massive halos or for a mass accretion history that varies with mass on cluster scales (see discussion in Chiu et al., 2016b). There is evidence for an increase in the fraction of cluster galaxies that are dominated by passively evolving stellar populations since $z \sim 1$ (H17), and this observed increase provides constraints on the timescales over which quenching of star formation occurs in those galaxies that are accreted by clusters (see, e.g., McGee et al., 2009).

Understanding the dynamics of galaxy accretion into clusters, from either lower mass clusters and groups or even individual systems from within the surrounding low density region, can shed additional light on galaxy evolution. A simulation based study argues that satellite orbits should become marginally more radial at higher redshifts, especially for systems with a higher host halo mass (Wetzel, 2011). Probes of redshift trends in the orbital characteristics of cluster galaxies have already been carried out (Biviano & Poggianti, 2009), providing some indication that passive galaxies have systematically different orbits at low and high redshift. In other studies of high redshift, relatively low mass systems, evidence has emerged that recently quenched galaxies have a preferred phase space

distribution that is different from that of passive galaxies (Muzzin et al., 2014; Noble et al., 2016).

In this paper, we attempt to build upon these studies by focusing on a dynamical analysis of galaxies within a large ensemble of Sunyaev-Zel'dovich effect (SZE) selected galaxy clusters extending to redshift $z \sim 1.3$. In contrast to these previous dynamical studies, our cluster sample has a well understood selection that does not depend on the galaxy properties, and the sample extends over a broad redshift range, allowing a cleaner examination of redshift trends. Moreover, each cluster has an SZE based mass estimate with ∼25 percent uncertainty (Bocquet et al., 2015), enabling us to estimate virial radii r_Δ with ~ 8 percent uncertainties and thereby ensuring that we are examining comparable regions of the cluster at all masses and redshifts.

Our goals are to study (1) whether there is evidence that the orbital characteristics of the passive galaxies are changing with redshift or mass in the cluster ensemble, (2) whether there is evidence within the galaxy dynamics for dynamical equilibrium and self-similarity with mass and redshift, and (3) whether the cluster mass constraints from our analysis are consistent with masses obtained through independent calibration in previously published SPT analyses.

Combining spectroscopic observations obtained at Gemini South, the VLT and the Magellan telescopes in a sample of 110 SPT-detected galaxy clusters, we construct a large sample of ∼3000 passive cluster members, spanning the wide redshift range of $0.2 < z < 1.3$. With this dataset we carry out a Jeans analysis (e.g. Binney & Tremaine, 1987) that adopts a framework of spherical symmetry and allows for a range of different velocity dispersion anisotropy profiles. Specifically, we use the Modeling Anisotropy and Mass Profiles of Observed Spherical Systems code (Mamon et al., 2013, hereafter MAMPOSSt) to explore the range of models consistent with the data, and then we use the results to characterize the velocity dispersion anisotropy profile, to test for evidence of virialization with the pseudo-phase-space density (PPSD) profile and to probe for trends with cluster mass or redshift in both. Exploring a broad range of possible velocity dispersion anisotropy profiles then allows us to extract robust constraints on the cluster virial masses as well. Throughout this paper, we address a number of limitations that have to be taken into account, such as the degeneracy between the mass and the velocity anisotropy profiles (see Section 3.2), the assumptions of spherical symmetry and dynamical equilibrium, and the presence of foreground/background interloper galaxies projected onto the cluster virial region. Mamon et al. (2013) have tested the accuracy of MAMPOSSt by analysing a sample of clusters extracted from numerical simulations, recovering r_{200} estimates with mean bias at $\leq 2.5\%$ and *rms* scatter of 6% for kinematic samples with 500 tracers.

The paper is organized as follows: In Section 3.2 we give an overview of the theoretical framework. In Section 3.3 we summarize the dataset used for our analysis. The results are presented in Section 3.4, where we discuss the outcome of our analysis of the velocity dispersion anisotropy profile, the PPSD profiles, the virial mass comparisons, and the impact of disturbed clusters on our analysis. We present our conclusions in Section 3.5. Throughout this paper we adopt a flat ΛCDM cosmology with the Hubble constant $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and assume the matter density parameter $\Omega_\mathrm{M} = 0.3$. The

virial quantities are computed at radius r_{200}. All quoted uncertainties are equivalent to Gaussian 1σ confidence regions, unless otherwise stated.

3.2 Theoretical Framework

The dynamical analysis implemented in this paper is based on the application of the Jeans equation to spherical systems (Binney & Tremaine, 1987). In spherical coordinates, the Jeans equation can be written as

$$\frac{d(\nu\sigma_r^2)}{dr} + \frac{\nu}{r}\left[2\sigma_r^2 - (\sigma_\theta^2 + \sigma_\phi^2)\right] = -\nu\frac{d\Phi}{dr}, \tag{3.1}$$

where ν is the number density profile of the tracer galaxy population, Φ is the gravitational potential, and $\sigma_r, \sigma_\theta, \sigma_\phi$ are the components of the velocity dispersion along the three spherical coordinates r, θ, ϕ. It is convenient to write this equation as

$$\frac{GM(r)}{r} = -\sigma_r^2\left(\frac{d\ln\nu}{d\ln r} + \frac{d\ln\sigma_r^2}{d\ln r} + 2\beta\right), \tag{3.2}$$

where $M(r)$ is the enclosed mass within radius r, G is Newton's constant, $\beta \equiv 1 - (\sigma_\theta^2/\sigma_r^2)$ is the velocity dispersion anisotropy that is generically a function of radius, and $\sigma_\theta = \sigma_\phi$.

In principle, it is therefore possible to use Eq. 3.2 to estimate the mass distribution $M(r)$ of the system. However, we as external observers can measure only projected quantities, such as the surface density profile of the tracer population $\Sigma(R)$ and the line-of-sight (LOS) velocity distribution (or, alternatively, the line-of-sight velocity dispersion $\sigma_{\rm LOS}(R)$). These two observed functions are not sufficient to derive a unique mass model— at least within the context of the typical observational uncertainties where full knowledge of the line of sight velocity distribution is lacking (Merritt, 1987). This degeneracy between the mass and the velocity anisotropy profiles can be addressed in several ways.

3.2.1 Dynamical analysis with MAMPOSSt

A first method consists of assuming that, at a given projected radius, the LOS velocity distribution can be described by a Gaussian (Strigari et al., 2008; Wolf et al., 2010). However, because the true distribution may deviate from Gaussianity— including the case of anisotropic systems (Merritt, 1987)— the constraints from this approach on the anisotropy are weak (Walker et al., 2009). A step forward, therefore, consists of analysing the kurtosis of the LOS velocity distribution. This was found to be a powerful tool to break the mass-anisotropy degeneracy (Merritt, 1987; Gerhard, 1993; van der Marel & Franx, 1993; Zabludoff et al., 1993; Łokas, 2002; Łokas & Mamon, 2003). Much effort has been put into further constraining the anisotropy taking into account more of the information contained in the projected phase-space velocity distribution as a function of projected radius, considering, for example, the full set of even moments (Kronawitter et al., 2000; Wojtak et al., 2008, 2009).

In this work, we fit the whole projected phase-space velocity distribution using the MAMPOSSt code (Mamon et al., 2013). We use this code to determine the mass and anisotropy profiles of a cluster in parametrized form by performing a likelihood exploration to the distribution of the cluster galaxies in projected phase-space, constraining the parameters describing these two profiles. This method is based on the assumption of spherical symmetry, and adopts a Gaussian as the shape of the distribution of the 3D velocities without demanding Gaussian LOS velocity distributions. We emphasize that it does not assume that light traces mass, allowing the scale radius r_s of the total mass distribution to differ from that of the galaxy distribution. For more details on the code, we refer the reader to Mamon et al. (2013).

MAMPOSSt requires parametrized models for the number density profile, the mass profile and the velocity anisotropy profile, without any limitation on the choice of these models. We will address the issue of the number density profile in Section 3.3.3.

The current estimates of the systematic error on MAMPOSSt derived dynamical masses are $\approx 10\%$, where this number comes from an analysis of clusters extracted from numerical simulations (see Mamon et al., 2013). In their study, they find that for the dynamical tracers defined to lie within a sphere of r_{100} that the estimate of the virial radius r_{200} is biased at $\leq 3.3\%$ (see Table 2; Mamon et al., 2013). Thus, as a prior on the virial mass M_{200} bias, we adopt a Gaussian with $\sigma = 10\%$ centered at no bias.

3.2.2 Mass and anisotropy profiles

For the mass profile in our analyses, we consider 5 models, namely the Navarro, Frenk and White profile (NFW; Navarro et al., 1996), the Einasto profile (Einasto, 1965), the Burkert profile (Burkert, 1995), the Hernquist profile (Hernquist, 1990), and the Softened Isothermal Sphere (SIS; Geller et al., 1999). All these models have been applied to galaxy clusters in previous works (e.g. Mohr et al., 1996; Rines et al., 2003; Katgert et al., 2004a; Rines & Diaferio, 2006; Biviano et al., 2006). As we show later in Section 3.4, our data cannot distinguish among these different mass profiles, and so in the analyses described below we adopt the NFW model

$$\rho(r) = \rho_0 \left(\frac{r}{r_s}\right)^{-1} \left(1 + \frac{r}{r_s}\right)^{-2} \tag{3.3}$$

where ρ_0 is the central density and r_s is the scale radius where the logarithmic derivative of the density profile reaches -2.

For the velocity anisotropy profile, we consider the following five models that have been used in previous MAMPOSSt analyses:

C a radially constant anisotropy model,

$$\beta_C(r) = \theta_\beta; \tag{3.4}$$

T the Tiret anisotropy profile (Tiret et al., 2007),

$$\beta_T(r) = \theta_\beta \frac{r}{r + r_s}, \tag{3.5}$$

which is isotropic at the center and characterized by the anisotropy value θ_β at large radii. The transition radius r_s is the scale radius of the NFW density profile;

O a model with anisotropy of opposite sign at the center and at large radii,

$$\beta_O(r) = \theta_\beta \frac{r - r_s}{r + r_s}; \tag{3.6}$$

MŁ the Mamon & Łokas (2005) profile,

$$\beta_{M\text{Ł}}(r) = 0.5 \frac{r}{r + \theta_\beta}; \tag{3.7}$$

OM the Osipkov-Merritt anisotropy profile (Osipkov, 1979; Merritt, 1985),

$$\beta_{OM}(r) = \frac{r^2}{r^2 + \theta_\beta^2}. \tag{3.8}$$

Summarizing, we run MAMPOSSt with 3 free parameters: the virial radius r_{200}, the scale radius r_s of the mass distribution, and a velocity anisotropy parameter θ_β. This parameter represents the usual $\beta = 1 - (\sigma_\theta^2/\sigma_r^2)$ for the first three models (C, T, O), while for the MŁ and OM models it defines a characteristic radius $\theta_\beta = r_\beta$. The maximum likelihood solutions are obtained using the NEWUOA software (Powell, 2006) and are shown in Section 3.4.

3.3 Cluster Data

In this section we present the cluster sample, spectroscopic data, and the method for constructing composite clusters.

3.3.1 Cluster sample

The cluster sample analyzed in this study consists of galaxy clusters detected with the South Pole Telescope (SPT), a 10-meter telescope located within 1 km of the geographical South Pole, observing in three mm-wave bands centered at 95, 150 and 220 GHz (see Carlstrom et al., 2011). The SPT-SZ survey, imaging 2500 deg^2 of the southern sky, has produced data that are used to select galaxy clusters via their thermal SZE signature in the 95 and 150 GHz maps. This SZE signature arises through the inverse Compton scattering of the Cosmic Microwave Background (CMB) photons and the hot intracluster medium (ICM; Sunyaev & Zeldovich, 1972). We refer the reader to Schaffer et al. (2011) for details on the survey strategy and data processing. All of the clusters studied in this work have a high detection significance ($\xi > 4.8$; Song et al., 2012b; Bleem et al., 2015) and have spectroscopic redshifts.

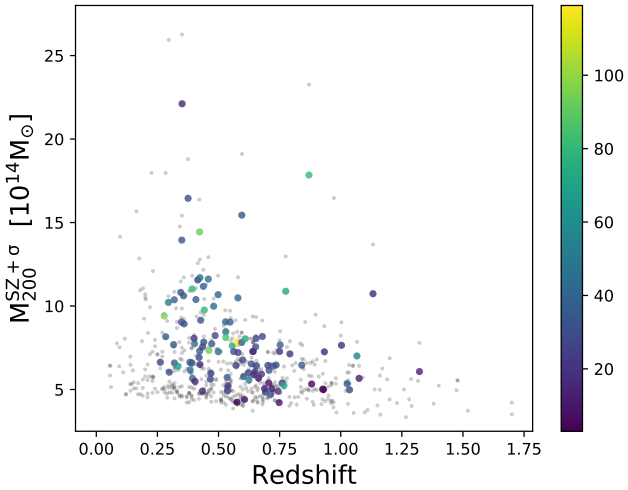

Figure 3.1: Masses vs redshifts of the cluster sample. The colored dots are coded according to the number of member galaxies in each cluster (see color bar on right). Overplotted in grey is the full distribution of the 2500 $\deg^2$ SPT-SZ sample.

A major advantage of selecting clusters with the SZE rather than other cluster observables lies in the fact that the surface brightness of the SZE signature is independent of the cluster redshift, which together with the expected temperature and density evolution of the cluster ICM at fixed virial mass and the changing solid angle of clusters with redshift lead an SZE signal to noise selected sample to be approximately mass-limited (Haiman et al., 2001; Holder et al., 2001).

Because we plan to carry out a dynamical analysis, the cluster sample we analyze includes only those SPT systems with spectroscopic follow-up. This spectroscopic subsample is not a signal to noise selected sample, but rather is the largest sample of SPT selected clusters we could assemble for the analysis. We plot the sample of 110 clusters in the space of redshift versus mass in Fig. 3.1, together with the full distribution of the 2500 $\deg^2$ SPT-SZ sample with the latest available redshifts (Bayliss et al., 2016; Khullar et al., 2018; Strazzullo et al., 2018, Bocquet et al., in prep). Note that in comparison to the SPT-SZ cosmology sample (de Haan et al., 2016), the median redshift and mass for our sample is 0.56 and $7.26 \times 10^{14} M_\odot$ as compared to 0.55 and $6.08 \times 10^{14} M_\odot$. Approximately 6 percent of the sample lies at $z > 1$ as compared to ~ 9 percent of the full SPT-SZ $\xi > 5$ sample (with updated redshifts from Bocquet et al., in prep). Thus, in comparison to the complete SPT-SZ sample, the subset of those clusters we study here have somewhat higher masses and are somewhat underrepresented at high redshift. Full information on the cluster sample is provided in Table 3.1. From left to right the columns correspond to the SPT cluster designation, the total number of passive galaxy redshifts available, the number of passive galaxy member redshifts, the cluster redshift, the SZE based mass $M_{200}^{\mathrm{SZ}+\sigma}$ and the corresponding virial radius $R_{200}^{\mathrm{SZ}+\sigma}$ (described in Section 3.3.3).

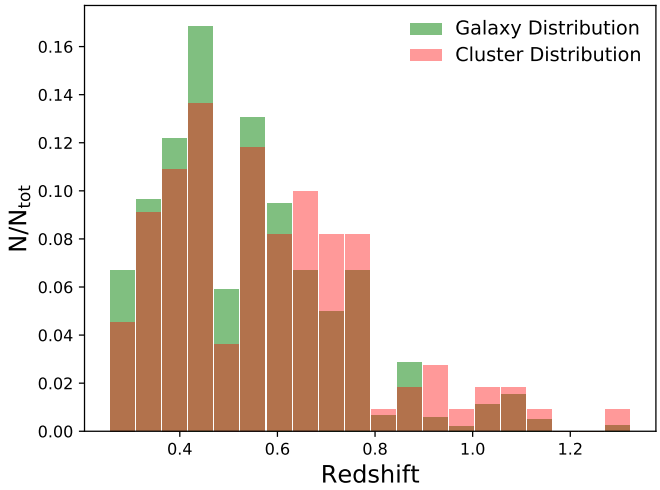

Figure 3.2: Redshift distribution of member galaxies (green) and clusters (red) in our sample, normalized to the total number of objects in each case. The two distributions are similar, but a clear trend to have fewer galaxies per cluster at high redshift is visible.

3.3.2 Spectroscopic sample

We use spectroscopic follow-up including data taken using the Gemini Multi Object Spectrograph (GMOS; Hook et al., 2004) on Gemini South, the Focal Reducer and low dispersion Spectrograph (FORS2; Appenzeller et al., 1998) on VLT Antu, the Inamori Magellan Areal Camera and Spectrograph (IMACS; Dressler et al., 2006) on Magellan Baade, and the Low Dispersion Survey Spectrograph (LDSS3; Allington-Smith et al., 1994) on Magellan Clay.

We combine the datasets presented in Ruel et al. (2014) and in Bayliss et al. (2016) with the one described in Sifón et al. (2013a), obtaining a sample of 4695 redshifts. In what follows, we split the spectroscopic galaxy sample according to the spectral features into two subsamples of those with emission lines (EL) and those without emission lines (nEL), adopting the determinations presented in the aforementioned analyses.

Due to the observational strategy of these surveys, which targeted red-sequence (RS) galaxies, the number of nEL galaxies (3834) greatly exceeds the numer of EL galaxies. Because EL and nEL galaxies are characterized by different spatial and kinematic properties (see, e.g., Mohr et al., 1996; Biviano et al., 1997; Dressler et al., 1999; Biviano et al., 2002, 2016; Bayliss et al., 2017; Hennig et al., 2017), in the analysis that follows we focus only on the passive nEL population, for which we have sufficient statistics to properly carry out the dynamical analysis. In Fig. 3.2 we show the normalized distribution of galaxies and clusters as a function of redshift.

Cluster	$N_{\rm tot}$	$N_{\rm mem}$	$z_{\rm clus}$	$M_{200}^{\rm SZ+\sigma}$ [$10^{14}{\rm M}_{\odot}$]	$R_{200}^{\rm SZ+\sigma}$ [Mpc]
SPT-CL J0000-5748	27	24	0.702	6.44	1.37
SPT-CL J0013-4906	37	36	0.408	10.39	1.81
SPT-CL J0014-4952	38	27	0.752	7.67	1.42
SPT-CL J0033-6326	28	14	0.599	6.77	1.45
SPT-CL J0037-5047	37	17	1.026	5.34	1.13
SPT-CL J0040-4407	33	33	0.350	13.95	2.04
SPT-CL J0102-4603	35	16	0.841	6.46	1.30
SPT-CL J0102-4915	81	80	0.870	17.84	1.80
SPT-CL J0106-5943	35	26	0.348	9.04	1.76
SPT-CL J0118-5156	8	7	0.705	5.40	1.29
SPT-CL J0123-4821	26	18	0.655	6.43	1.40
SPT-CL J0142-5032	28	23	0.679	8.17	1.50
SPT-CL J0200-4852	45	34	0.499	7.26	1.55
SPT-CL J0205-5829	15	8	1.322	6.07	1.06
SPT-CL J0205-6432	19	12	0.744	4.90	1.23
SPT-CL J0212-4657	26	20	0.654	8.06	1.51
SPT-CL J0232-5257	77	61	0.556	7.61	1.54
SPT-CL J0233-5819	9	9	0.664	5.68	1.33
SPT-CL J0234-5831	24	21	0.415	11.56	1.87
SPT-CL J0235-5121	96	82	0.278	9.41	1.84
SPT-CL J0236-4938	66	63	0.334	6.39	1.58
SPT-CL J0240-5946	19	17	0.400	8.06	1.67
SPT-CL J0243-4833	39	37	0.498	10.68	1.76
SPT-CL J0243-5930	32	25	0.634	6.53	1.42
SPT-CL J0252-4824	27	22	0.421	6.92	1.57
SPT-CL J0254-5857	37	32	0.438	11.18	1.83
SPT-CL J0257-5732	15	14	0.434	4.91	1.39
SPT-CL J0304-4401	45	35	0.458	11.62	1.84
SPT-CL J0304-4921	79	72	0.392	11.02	1.85
SPT-CL J0307-6225	26	17	0.580	7.21	1.49
SPT-CL J0310-4647	33	28	0.707	6.15	1.35
SPT-CL J0317-5935	25	18	0.469	5.97	1.47
SPT-CL J0324-6236	19	9	0.750	7.28	1.40
SPT-CL J0330-5228	80	71	0.442	9.77	1.75
SPT-CL J0334-4659	29	25	0.486	8.23	1.62
SPT-CL J0346-5439	85	79	0.530	8.11	1.59
SPT-CL J0348-4515	31	24	0.359	8.93	1.75
SPT-CL J0352-5647	22	16	0.649	6.07	1.37
SPT-CL J0356-5337	26	5	1.036	4.98	1.10
SPT-CL J0403-5719	31	24	0.467	5.62	1.44
SPT-CL J0406-4805	28	26	0.736	6.46	1.35
SPT-CL J0411-4819	45	42	0.424	11.70	1.87
SPT-CL J0417-4748	40	30	0.579	10.49	1.69
SPT-CL J0426-5455	15	11	0.642	7.28	1.46
SPT-CL J0433-5630	24	18	0.692	4.84	1.25
SPT-CL J0438-5419	92	87	0.422	14.44	2.00
SPT-CL J0449-4901	20	16	0.792	7.13	1.37
SPT-CL J0456-5116	31	20	0.562	7.27	1.51
SPT-CL J0509-5342	93	88	0.461	7.34	1.58
SPT-CL J0511-5154	18	14	0.645	5.90	1.36
SPT-CL J0516-5430	51	47	0.295	10.22	1.88
SPT-CL J0521-5104	21	21	0.675	5.92	1.35
SPT-CL J0528-5300	75	63	0.768	5.24	1.25
SPT-CL J0533-5005	4	4	0.881	5.33	1.20
SPT-CL J0534-5937	3	3	0.576	4.24	1.25
SPT-CL J0540-5744	24	17	0.760	5.38	1.26
SPT-CL J0542-4100	36	29	0.640	7.27	1.46
SPT-CL J0546-5345	54	49	1.066	7.01	1.22
SPT-CL J0549-6205	31	26	0.375	16.45	2.13
SPT-CL J0551-5709	39	30	0.423	7.30	1.60

Table 3.1: Cluster spectroscopic sample: We present the cluster name, total number of passive galaxies in the spectroscopic sample $N_{\rm tot}$, members used in our analysis after interloper rejection $N_{\rm mem}$, cluster redshift $z_{\rm clus}$, and cluster mass $M_{200}^{\rm SZ+\sigma}$ and corresponding virial radius $R_{200}^{\rm SZ+\sigma}$ from Bocquet et al. (2015).

Cluster	$N_{\rm tot}$	$N_{\rm mem}$	$z_{\rm clus}$	$M_{200}^{\rm SZ+\sigma}$ $[10^{14}{\rm M}_{\odot}]$	$R_{200}^{\rm SZ+\sigma}$ [Mpc]
SPT-CL J0555-6406	34	27	0.345	10.82	1.88
SPT-CL J0559-5249	72	67	0.609	8.04	1.53
SPT-CL J0655-5234	33	30	0.472	7.74	1.60
SPT-CL J2017-6258	38	35	0.535	5.73	1.41
SPT-CL J2020-6314	27	18	0.537	4.91	1.34
SPT-CL J2022-6323	34	29	0.383	6.63	1.57
SPT-CL J2026-4513	15	11	0.688	5.05	1.27
SPT-CL J2030-5638	42	36	0.394	5.67	1.48
SPT-CL J2032-5627	39	32	0.284	8.16	1.75
SPT-CL J2035-5251	42	29	0.529	9.05	1.65
SPT-CL J2040-5725	7	4	0.930	5.00	1.15
SPT-CL J2043-5035	33	21	0.723	6.40	1.36
SPT-CL J2056-5459	13	11	0.718	5.06	1.26
SPT-CL J2058-5608	10	6	0.607	4.40	1.25
SPT-CL J2100-4548	37	18	0.712	4.68	1.23
SPT-CL J2106-5844	16	16	1.132	10.74	1.37
SPT-CL J2115-4659	31	28	0.299	6.03	1.57
SPT-CL J2118-5055	55	33	0.624	5.57	1.35
SPT-CL J2124-6124	23	21	0.435	7.56	1.61
SPT-CL J2130-6458	41	40	0.316	7.69	1.69
SPT-CL J2135-5726	31	30	0.427	9.15	1.72
SPT-CL J2136-4704	19	19	0.424	6.50	1.53
SPT-CL J2136-6307	6	6	0.926	5.02	1.15
SPT-CL J2138-6008	39	32	0.319	10.37	1.87
SPT-CL J2140-5727	17	11	0.404	5.49	1.46
SPT-CL J2145-5644	43	35	0.480	9.99	1.73
SPT-CL J2146-4633	15	7	0.933	7.26	1.30
SPT-CL J2146-4846	26	25	0.623	6.00	1.38
SPT-CL J2146-5736	34	23	0.602	5.77	1.38
SPT-CL J2148-6116	24	24	0.571	6.45	1.45
SPT-CL J2155-6048	23	19	0.539	5.23	1.36
SPT-CL J2159-6244	38	36	0.391	6.59	1.56
SPT-CL J2222-4834	29	25	0.652	7.56	1.48
SPT-CL J2232-5959	34	26	0.595	7.82	1.53
SPT-CL J2233-5339	33	28	0.440	7.84	1.62
SPT-CL J2248-4431	15	14	0.351	22.11	2.37
SPT-CL J2300-5331	25	21	0.262	6.63	1.64
SPT-CL J2301-5546	12	8	0.748	4.21	1.17
SPT-CL J2306-6505	46	42	0.530	8.46	1.61
SPT-CL J2325-4111	33	27	0.357	10.61	1.86
SPT-CL J2331-5051	119	108	0.576	7.85	1.54
SPT-CL J2332-5358	47	45	0.402	7.74	1.64
SPT-CL J2335-4544	37	33	0.547	9.04	1.63
SPT-CL J2337-5942	28	19	0.775	10.88	1.59
SPT-CL J2341-5119	18	13	1.003	7.65	1.29
SPT-CL J2342-5411	12	7	1.075	5.67	1.14
SPT-CL J2344-4243	33	25	0.595	15.44	1.91
SPT-CL J2351-5452	42	30	0.384	6.18	1.53
SPT-CL J2355-5055	36	33	0.320	6.59	1.61
SPT-CL J2359-5009	44	37	0.775	5.17	1.24

Table 3.1: Cluster spectroscopic sample - Continued.

Bin	Redshift range	$\langle z \rangle$	$N_{\rm clus}$	$N_{\rm tot}$	$N_{\rm mem}$	$\langle M_{200}^{\rm SZ+\sigma} \rangle$ $[10^{14} M_{\odot}]$
1	0.26-0.38	0.33	18	712	593	9.23
2	0.39-0.44	0.42	19	744	644	9.50
3	0.46-0.56	0.51	17	758	615	7.98
4	0.56-0.71	0.62	30	904	675	7.31
5	0.71-1.32	0.86	26	716	459	7.60
-	0.26-1.32	0.55	110	3834	2966	8.28

Table 3.2: Characteristics of the composite clusters in redshift: columns show the bin number, redshift range, mean redshift, number of included clusters $N_{\rm clus}$, number of total galaxies $N_{\rm tot}$, number of member galaxies $N_{\rm mem}$ and mean SPT based mass $\langle M_{200}^{\rm SZ+\sigma} \rangle$.

3.3.3 Construction of composite clusters

To enable a precise determination of the cluster masses and the velocity anisotropy profiles, we cannot rely on the few spectroscopic members in the individual clusters of our sample. For this reason, we either create composite clusters with much more dynamical information or, as for the results presented in Section 3.4.4, fit to a common model across the cluster ensemble by combining the likelihoods associated with each individual cluster. The composite cluster approach has been adopted in previous analyses (Carlberg et al., 1997b; van der Marel et al., 2000a; Katgert et al., 2004a; Biviano & Poggianti, 2009), and it is supported by cosmological simulations that predict cosmological halos can be characterized by a universal mass density profile with a concentration that depends mildly on the halo mass (Navarro et al., 1997). We return to the method using a combination of individual cluster likelihoods in Section 3.4.4.

We create composite clusters by combining cluster subsamples within different mass and redshift ranges in such a way that each subsample includes an adequate number of members. Considering that tests done using MAMPOSSt on cosmological simulations indicate that with 500 tracers the code is able to recover the mass and anisotropy parameters with a suitably small uncertainty (Mamon et al., 2013), we create similarly sized subsamples. With this approach we can construct 5 composite clusters selected either as redshift or mass subsamples.

Table 3.2 lists the characteristics of the composite clusters created within redshift bins. From left to right the columns correspond to the bin number, the redshift range of the clusters combined to create the composite system in that bin, the mean redshift of these clusters, the number of clusters, the total number of galaxy redshifts, the number of cluster members with redshifts, and the average SPT based mass of the individual clusters in the sample. In the following subsections we describe in detail how these composite clusters are created and how interlopers are removed from the spectroscopic samples.

Rescaling observables with SZE based mass estimates

To create composite clusters we choose the Brightest Cluster Galaxy (BCG) as the cluster center. We determine these positions using results from previous and ongoing work (Song et al., 2012b; Sifón et al., 2013a, Stalder et al., in prep), allowing us to calculate a projected separation R for each galaxy from the cluster center. We extract the rest frame LOS velocity $\mathrm{v_{rf}}$ from the galaxy redshift z and equivalent velocity $\mathrm{v}(z)$ as $\mathrm{v_{rf}} \equiv (\mathrm{v}(z) - \mathrm{v}(z_\mathrm{c}))/(1+z_\mathrm{c})$, where z_c is the cluster redshift.

Because our clusters span a wide range of mass and redshift, we must scale the galaxy projected cluster distances R and LOS velocities to the values of a fiducial cluster mass using an estimate of the individual cluster virial radius and velocity. For convenience, in the composite cluster we use the mean mass $\langle M_{200} \rangle$ and the mean redshift $\langle z \rangle$ of the ensemble of clusters in that bin to calculate the associated $\langle R_{200} \rangle$ and $\langle \mathrm{v_{200}} \rangle$ that we adopt as fiducial values for the bin. Thus, the rescaled observables for each galaxy i within a cluster j in a particular bin are $R_\mathrm{i,j} = R_\mathrm{i} \langle R_{200} \rangle / R_\mathrm{200,j}$ and $\mathrm{v_{i,j}} = \mathrm{v_{rf,i}} \langle \mathrm{v_{200}} \rangle / \mathrm{v_{200,j}}$, where $R_\mathrm{200,j}$ and $\mathrm{v_{200,j}}$ are the virial radius and circular velocity, respectively, of cluster j. The circular velocity is defined using the virial condition $\mathrm{v}_{200}^2 = GM_{200}/R_{200}$.

To perform the normalization, we need precise estimates of the cluster mass. Given that the SPT sample is SZE selected, we adopt the SZE observable ξ, which is the detection signal to noise and is correlated to the underlying cluster virial mass (Andersson et al., 2011; Benson et al., 2013), as the source of our cluster mass estimates. SPT masses derived from ξ have a statistical uncertainty that depends on the intrinsic scatter in the ξ–mass relation and on the observational uncertainty in the signal to noise of the detection; together, these lead characteristically to $\sim$20 percent statistical uncertainty in the cluster virial mass. In addition, there are systematic uncertainties remaining from the calibration of the mass–observable relation that are currently at the $\sim$15 percent level. As discussed further below, the systematic mass scale uncertainties at this level have no impact on our analysis, but the cluster to cluster statistical mass uncertainty of $\sim$20 percent introduces a corresponding uncertainty in R_{200} and v_{200} of $\sim$7 percent.

The mass–observable relation for the SPT sample can be calibrated in different ways, and in the current analysis we use two different approaches. The first approach uses direct cluster mass measurements from weak lensing, velocity dispersions, X-ray measurements to calibrate the SZE-mass observable relation. The second approach fits the SPT cluster distribution in signal to noise ξ and redshift z and the mass-observable relation to a flat ΛCDM model with external cosmological priors from other datasets. These constraints from, e.g., CMB anisotropy, baryon acoustic oscillations, etc., effectively constrain the mass-observable relation so that the cluster mass function implied by SPT cluster distribution in ξ and z is consistent with that expected given the external cosmological priors. As an example (see Bocquet et al., 2015), the inclusion of external cosmological constraints from Planck CMB in the SPT mass–observable calibration leads to a shift of cluster masses to $\sim$25 percent higher values in comparison to the case where the mass calibration is undertaken with direct cluster mass measurements from velocity dispersions.

Given the need to have an initial mass estimate to enable the stacking of the clusters for

a dynamical analysis, we carry out the analysis with initial SPT based masses calibrated in two different ways. The first uses the SPT cluster counts together with direct cluster mass information from velocity dispersions. Here we refer to these masses as $M_{200}^{\mathrm{SZ}+\sigma}$. The second uses the SPT cluster counts together with external cosmological priors from Planck CMB anisotropy and distance measurements (BAO, Supernovae), specifically we use mass-calibration inferred from the SPTCL +Planck +WP+BAO+SNIa data set in Bocquet et al. (2015). We refer to these masses as $M_{200}^{\mathrm{SZ+Planck}}$. In both cases the SPT mass–observable relation is calibrated using M_{500} (Bocquet et al., 2015), and so we transform from M_{500} by assuming an NFW model with a concentration c_{200} sampled from cosmological N-body simulations (Duffy et al., 2008). The r_{200} and v_{200} are then easily obtained, depending only on the cluster redshift and the adopted cosmology.

As discussed in Section 3.4.4, the dynamical mass measurements of the composite clusters have only a weak dependence on the mass scale of the velocities and radii used to build the composite clusters; specifically, the ∼25 percent shift between the masses $M_{200}^{\mathrm{SZ}+\sigma}$ and $M_{200}^{\mathrm{SZ+Planck}}$ used to build the two sets of composite clusters has no significant impact on the final dynamical masses.

Because we are carrying out a Jeans analysis, which is based on the assumption of dynamical equilibrium, we restrict our analysis to the cluster virial region ($R \leq r_{200}$). Moreover, we exclude the very central cluster region ($R \leq 50$ kpc). In fact, the composite clusters are centered on the BCG, which we exclude from the dynamical analysis. We note also that in the composite clusters the characteristic asphericity of individual clusters is averaged down, leading to a combined system that is approximately spherical in agreement with the dynamical model we are employing.

Interloper rejection

One benefit of constructing composite clusters is that we can more easily identify and reject some interloper galaxies, i.e., galaxies that are projected inside the cluster virial region, but do not actually lie inside it. We do so by using the "Clean" method (Mamon et al., 2013), which is based on the identification of the cluster members on the basis of their projected phase-space location (R, v). The LOS velocity dispersion σ_{LOS} of the composite cluster is used to estimate the cluster mass using a scaling relation calibrated using numerical simulations (e.g., Saro et al., 2013), and an NFW mass profile with concentration sampled from the theoretical mass-concentration relation of Macciò et al. (2008a). Thereafter, assuming the MŁ velocity anisotropy profile model, and given the $M(r)$ of the cluster, an LOS velocity dispersion profile $\sigma_{\mathrm{LOS}}(R)$ is calculated and used to iteratively reject galaxies with $|\mathrm{v}_{\mathrm{rf}}| > 2.7\sigma_{\mathrm{LOS}}$ at any clustercentric distance (see Mamon et al., 2010b, 2013). As an example, we present in Fig. 3.3 the location of galaxies in projected phase-space with the identification of cluster member galaxies for the full sample.

We note that even after cleaning the dynamical sample, it is still contaminated to some degree by interlopers. One reason for this is that galaxies near the cluster turn-around radius (i.e., well outside the virial radius) will have small LOS velocities. Therefore, if those galaxies are projected onto the cluster virial region they simply cannot be separated

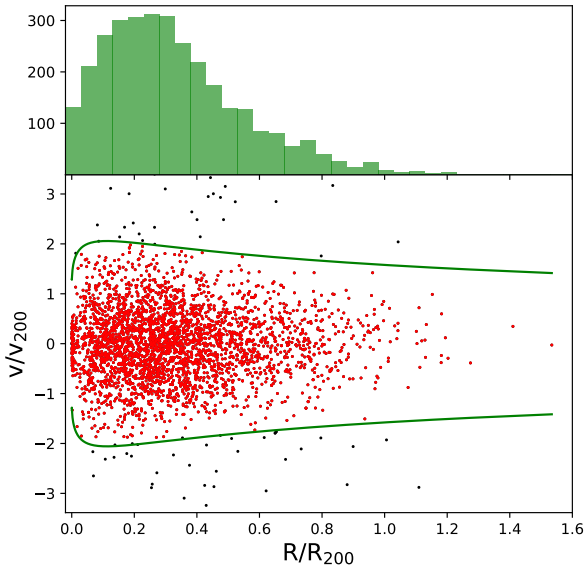

Figure 3.3: The projected phase-space diagram (below) in the rescaled coordinates $(R/R_{200}, \mathrm{v}/\mathrm{v}_{200})$ for the full sample. Green lines represent the radially-dependent $2.7\sigma_{\mathrm{LOS}}$ cut used to reject interlopers (indicated by black dots). The radial distribution of the member galaxies with measured redshifts is shown in the upper panel.

from the galaxies that actually lie within the cluster virial radius. Analysis of cosmological N-body simulations shows that when passive galaxies are selected this contamination is characteristically ~ 20 percent (Saro et al., 2013) for SPT mass scale clusters and represents a systematic, whose effects on the dynamical analysis require further exploration (see also discussion in Old et al., 2015).

Galaxy number density profiles

We cannot simply adopt the spectroscopic sample to measure the number density profiles of the galaxy population, because the structure of the spectrographs and our observing strategy of visiting each cluster typically with only two masks will generally result in a radially dependent incompleteness. This incompleteness, if not accounted for, would affect the determination of the cluster projected number density profile $\Sigma(R)$, related to the 3D number density profile $\nu(r)$ through the Abel inversion (Binney & Tremaine, 1987). As only the logarithmic derivative of $\nu(r)$ enters the Jeans equation (see equation (3.2)), the absolute normalization of the galaxy number density profile has no impact on our analysis. However, a radially dependent incompleteness in the velocity sample would impact our analysis.

Thus, for our analysis we rely on a study of the galaxy populations in 74 SZE selected clusters from the SPT survey that were imaged as part of the Dark Energy Survey (DES) Science Verification phase (H17). This complete sky-area-selected subsample of the SPT-SZ cluster sample has a redshift and mass distribution that are consistent with our sample. The

$\beta(r)$ model	Mass model				
	NFW	Burkert	Einasto	Hernquist	SIS
C	0.52	0.19	0.47	1.00	9.04×10^{-4}
ML	0.68	0.07	1.00	0.89	7.46×10^{-4}
OM	0.45	0.18	0.46	1.00	4.44×10^{-6}
O	0.90	0.03	1.00	0.43	
T	0.75	0.06	0.84	1.00	

Table 3.3: Likelihood ratio (or Bayes factor; see equation 3.11) for the NFW, Burkert, Einasto, Hernquist and SIS mass profiles is presented for each anisotropy profile analyzed using the composite cluster constructed using the full dynamical sample. As noted in the text, we cannot discard any of the mass models aside from the SIS.

number density profile of the red sequence population of galaxies is found to be well fit by an NFW model out to radii of $4r_{200}$, with a concentration of cluster galaxies $c_{\rm gal} = 5.37^{+0.27}_{-0.24}$. The non-red sequence population is much less centrally concentrated, but in our analysis we are focused only on the nEL galaxy populations. No statistically significant redshift or mass trends in the galaxy concentration were identified. Therefore, we adopt the number density profile described by an NFW profile with the above-mentioned value of $c_{\rm gal}$ and a scale radius $r_{\rm s,\,gal} = R_{200}/c_{\rm gal}$. Implicit in this approach is the assumption that the dynamical properties of our spectroscopic sample are consistent with the dynamical properties of the red sequence galaxy population used to measure the radial profiles.

We test this assumption of a fixed $c_{\rm gal}$ by including in the Likelihood an additional parameter $r_{\rm s,\,gal}$, given by the ratio $R_{200}/c_{\rm gal}$, where $c_{\rm gal}$ is given by H17

$$c_{\rm gal} = A \left(\frac{M_{200}}{M_{piv}} \right)^{B} \left(\frac{1+z}{1+z_{piv}} \right)^{C}, \tag{3.9}$$

where, for red sequence galaxies, $A = 5.47 \pm 0.53, B = -0.01 \pm 0.10, C = 0.15 \pm 0.30$. We choose mass and redshift pivot points $M_{piv} = 5 \times 10^{14}$ and $z_{piv} = 0.53$, respectively, corresponding to the median mass and redshift of our sample. With this additional parameter, fitting for A, B, and C with Gaussian priors corresponding to the uncertainties listed above, we adopt the Markov Chain Monte Carlo (MCMC) method for sampling from the probability distribution, utilizing the *emcee* code (Foreman-Mackey et al., 2013). Our analysis provides values of the scale radius and virial radius consistent within 1σ of the measured values reported in H17.

3.4 Results

In this section we present the results of our dynamical analysis. In the first subsection we explore the mass profiles and present several crucial pieces of information that are needed to understand the results in the following sections. The following subsections present our

$\beta(r)$ model	r_{200} [Mpc]	$r_{\rm s}$ [Mpc]	θ_β	Bayes Factor
	$\langle z\rangle = 0.33$; $r_\nu = 0.33$ Mpc			
C	$1.84^{+0.12}_{-0.07}$	$0.34^{+0.54}_{-0.12}$	$0.28^{+0.91}_{-0.24}$	0.39
MŁ	$1.79^{+0.14}_{-0.05}$	$0.31^{+0.25}_{-0.08}$	$0.37^{+3.99}_{-0.37}$	0.94
OM	$1.68^{+0.17}_{-0.07}$	$0.19^{+0.13}_{-0.05}$	$1.97^{+5.19}_{-0.77}$	1.00
O	$1.87^{+0.17}_{-0.14}$	$0.24^{+0.15}_{-0.06}$	$-0.56^{+2.38}_{-0.43}$	0.15
T	$1.81^{+0.09}_{-0.09}$	$0.27^{+0.19}_{-0.08}$	$0.19^{+2.08}_{-0.38}$	0.28
	$\langle z\rangle = 0.42$; $r_\nu = 0.32$ Mpc			
C	$1.85^{+0.10}_{-0.07}$	$0.44^{+0.53}_{-0.14}$	$0.20^{+0.90}_{-0.27}$	0.09
MŁ	$1.79^{+0.12}_{-0.06}$	$0.40^{+0.17}_{-0.11}$	$0.86^{+8.79}_{-0.56}$	1.00
OM	$1.62^{+0.17}_{-0.05}$	$0.27^{+0.15}_{-0.06}$	$1.19^{+2.97}_{-0.37}$	0.25
O	$1.78^{+0.15}_{-0.10}$	$0.35^{+0.12}_{-0.07}$	$0.39^{+1.66}_{-0.41}$	0.12
T	$1.81^{+0.09}_{-0.09}$	$0.42^{+0.22}_{-0.10}$	$0.39^{+1.59}_{-0.32}$	0.18
	$\langle z\rangle = 0.51$; $r_\nu = 0.29$ Mpc			
C	$1.69^{+0.08}_{-0.05}$	$0.32^{+0.23}_{-0.09}$	$0.19^{+0.77}_{-0.26}$	0.02
MŁ	$1.66^{+0.08}_{-0.05}$	$0.33^{+0.13}_{-0.10}$	$0.53^{+5.13}_{-0.30}$	0.32
OM	$1.51^{+0.07}_{-0.04}$	$0.33^{+0.19}_{-0.12}$	$0.76^{+1.35}_{-0.20}$	0.13
O	$1.56^{+0.06}_{-0.05}$	$0.31^{+0.07}_{-0.05}$	$0.93^{+0.21}_{-0.05}$	1.00
T	$1.64^{+0.06}_{-0.06}$	$0.40^{+0.17}_{-0.09}$	$0.68^{+0.88}_{-0.18}$	0.21
	$\langle z\rangle = 0.62$; $r_\nu = 0.27$ Mpc			
C	$1.53^{+0.06}_{-0.05}$	$0.35^{+0.31}_{-0.11}$	$0.39^{+0.64}_{-0.20}$	0.27
MŁ	$1.52^{+0.05}_{-0.05}$	$0.39^{+0.13}_{-0.13}$	$0.06^{+0.50}_{-0.05}$	0.15
OM	$1.43^{+0.09}_{-0.04}$	$0.18^{+0.09}_{-0.04}$	$2.40^{+5.74}_{-0.81}$	1.00
O	$1.51^{+0.06}_{-0.06}$	$0.23^{+0.09}_{-0.05}$	$-0.02^{+1.77}_{-0.35}$	0.09
T	$1.51^{+0.05}_{-0.06}$	$0.27^{+0.15}_{-0.07}$	$0.34^{+1.46}_{-0.30}$	0.22
	$\langle z\rangle = 0.86$; $r_\nu = 0.25$ Mpc			
C	$1.30^{+0.06}_{-0.05}$	$0.26^{+0.31}_{-0.09}$	$0.23^{+1.04}_{-0.31}$	0.12
MŁ	$1.28^{+0.07}_{-0.04}$	$0.25^{+0.12}_{-0.08}$	$0.58^{+7.20}_{-0.38}$	1.00
OM	$1.19^{+0.09}_{-0.03}$	$0.19^{+0.11}_{-0.05}$	$1.08^{+3.07}_{-0.44}$	0.32
O	$1.25^{+0.06}_{-0.05}$	$0.25^{+0.08}_{-0.05}$	$0.68^{+1.01}_{-0.24}$	0.33
T	$1.28^{+0.05}_{-0.06}$	$0.30^{+0.17}_{-0.08}$	$0.55^{+1.36}_{-0.26}$	0.36

Table 3.4: Parameter constraints from the MAMPOSSt analysis of the composite clusters defined in Table 3.2. Columns represent the velocity anisotropy model $\beta(r)$, the virial radius r_{200}, the scale radius $r_{\rm s}$, the anisotropy parameter θ_β and the Bayes factor from equation 3.11. For each composite cluster, we also report the value of r_ν, which is the scale radius of the galaxy number density profile, obtained from the ratio between the mean value of the $R_{200}^{\rm SZ+\sigma}$ in each redshift bin, and the fixed c_{gal} value (see Section 3.3.3). Note that for the anisotropy models C, T, and O the anisotropy parameter does not have units, while for the MŁ and OM the values are evaluated in Mpc.

Bin	Redshift range	$R_{200}^{\rm dyn}$ [Mpc]	$r_{\rm s}$ [Mpc]	$M_{200}^{\rm dyn}$ [$10^{14} M_{\odot}$]	c_{200}
1	0.26-0.38	$1.81^{+0.11}_{-0.10}$	$0.26^{+0.16}_{-0.10}$	$9.44^{+1.70}_{-1.65}$	$5.4^{+2.6}_{-2.1}$
2	0.39-0.44	$1.82^{+0.10}_{-0.09}$	$0.36^{+0.17}_{-0.11}$	$10.57^{+1.93}_{-1.55}$	$4.1^{+1.5}_{-1.3}$
3	0.46-0.56	$1.56^{+0.11}_{-0.06}$	$0.31^{+0.15}_{-0.13}$	$7.42^{+1.58}_{-0.92}$	$4.4^{+1.7}_{-1.8}$
4	0.56-0.71	$1.51^{+0.05}_{-0.06}$	$0.25^{+0.11}_{-0.08}$	$7.31^{+1.13}_{-0.62}$	$5.2^{+1.9}_{-1.9}$
5	0.71-1.32	$1.28^{+0.06}_{-0.06}$	$0.24^{+0.11}_{-0.10}$	$6.20^{+0.85}_{-0.88}$	$4.2^{+2.0}_{-1.6}$
-	0.26-1.32	$1.62^{+0.03}_{-0.05}$	$0.29^{+0.06}_{-0.07}$	$8.71^{+0.52}_{-0.80}$	$5.1^{+1.1}_{-1.0}$

Table 3.5: Parameter constraints for the composite cluster mass profiles. Columns represent the redshift range, virial radius $R_{200}^{\rm dyn}$, scale radius $r_{\rm s}$, dynamical mass $M_{200}^{\rm dyn}$ and concentration. These results are obtained using Bayesian model averaging of the different anisotropy models.

measurements of the anisotropy profiles, the pseudo phase-space density profiles, and the dynamical mass constraints. The final subsection examines the impact on our analysis of ongoing mergers.

3.4.1 Testing Mass and Anisotropy Profiles

As a first step, we use MAMPOSSt to analyze the composite cluster constructed using the full dynamical dataset on all mass and anisotropy models described in Section 3.2.1. Our goal is to determine which mass models are appropriate for this study. For each mass model, we explore all the different anisotropy models.

Comparison of different mass profiles

To quantitatively differentiate among the mass models (NFW, Einasto, Burkert, Hernquist and SIS), we compare the likelihood of the data being consistent with the model, employing the so-called Bayes factor. This factor is the marginalized likelihood of the model (see Hoeting et al., 1999, and references therein). It is computed by averaging the likelihood in a specific model M_j over the available prior range $P(\theta_j\,|M_j)$, reading

$$\mathcal{L}(D\,|M_j) = \int \mathcal{L}(D\,|\theta_j, M_j) P(\theta_j\,|M_j) \mathrm{d}\theta_j, \tag{3.10}$$

where $\mathcal{L}(D\,|\theta_j, M_j)$ is the likelihood of the data D given the model parameters θ_j.

Considering J models M_1, ..., M_J, we define the Bayes factor $B_{j,max}$ of each model j by normalizing by the most probable model, yielding

$$B_{j,max} = \frac{\mathcal{L}(D\,|M_j)}{\mathcal{L}(D\,|M_{\rm max})}, \tag{3.11}$$

where $M_{\rm max}$ indicates the model with the highest marginalized likelihood. The averaged posterior distribution on the parameters common to all models is then simply given by

the weighted average of the posterior distributions of each model, with the Bayes factor as weight. Models can also be rejected using their Bayes factors. According to the Jeffreys scale (Jeffreys, 1998), M_j is considered decisively rejectable if $B_{j,\max} < 0.01$.

Table 3.3 contains the measured values. One can see that for each anisotropy model there is one preferred model (Bayes factor 1.0), but that the likelihood ratios for all but one of the mass profile models are on the order of 1. Note that for the SIS profile we can only consider three of the five anisotropy models, because the T and O models need the value of the scale radius in the density profile, and that is not uniquely defined for the SIS. Indeed, with the exception of the SIS model, we cannot reject any of the mass models we consider here.

We note that our choice of parameter priors for these analyses does not affect the calculation of the Bayes factor, because these priors are set to be flat with allowed ranges that extend beyond the preferred range of each parameter.

Taking guidance from both theoretical expectation and observational results, we select the NFW model as a good description of the mass profile. As mentioned in Section 3.1, in fact, cosmological simulations produce DM halos with mass profiles well described by an NFW profile. This result is in good agreement with a variety of observational analyses using both dynamics and weak lensing (Carlberg et al., 1997b; van der Marel et al., 2000a; Biviano & Girardi, 2003; Katgert et al., 2004a; Umetsu et al., 2011), even though some results have preferred different models (Merritt et al., 2006; Navarro et al., 2010; Dutton & Macciò, 2014; van der Burg et al., 2015a; Sereno & Ettori, 2017).

Bayesian model averaging with anisotropy models

In contrast to the mass profiles, the literature does not provide us with strong predictions for the radial form of the velocity anisotropy or β profile. In Table 3.4 we list– for each of the composite clusters as defined in Table 3.2– the results of the MAMPOSSt analysis, with the anisotropy parameter θ_β being β_0 for the C model, β_∞ for the T and O models and r_β for the MŁ and OM models. The errors on each of the parameters listed in the table are obtained by a marginalization procedure, i.e. by integrating the probabilities $p(r_{200}, r_s, \beta)$ provided by MAMPOSSt over the remaining two free parameters. In addition to the anisotropy parameters, Table 3.4 contains the dynamical constraints on the composite cluster virial radius $R_{200}^{\rm dyn}$ and the Bayesian weight described above.

Because we cannot strongly reject any of the models, we combine the results from the different anisotropy profiles by performing a Bayesian model averaging, weighting every model by its Bayes factor and combining statistics from the different β models. This approach was first proposed by Vázquez et al. (2012a), and has subsequently been used in multiple analyses aside from our own (Vázquez et al., 2012b; Aslanyan et al., 2014; Hee et al., 2016; Planck Collaboration et al., 2016b). For a proof that the weighted and coadded probability distribution functions consistute a proper probability distribution function, we direct the reader to examine section 8.2, equation (69) of the last reference. The results of this analysis are shown in Table 3.5 where we present the virial radius $R_{200}^{\rm dyn}$, the NFW scale radius r_s, the virial mass $M_{200}^{\rm dyn}$, and the concentration c_{200}. The 1σ parameter uncertainties

are computed through a marginalization procedure, as before.

Impact of Mass Priors on Composite Clusters

As mentioned in Section 3.3.3, when creating the composite clusters, we perform a rescaling of the observables (the galaxy projected cluster distances R and rest frame velocities $v_{\rm rf}$) using two different initial mass estimates, $M_{200}^{\rm SZ+\sigma}$ and $M_{200}^{\rm SZ+Planck}$ (available for each individual cluster). The dynamical mass constraints that result from the composite clusters created using these two different scalings do not differ significantly, demonstrating the stability of our analysis to underlying systematic mass uncertainties in the initial mass estimates used to construct the composite clusters. We return to this point in Section 3.4.4 below, where we show the dynamical masses obtained in the two cases (see Table 3.6).

Moreover, the masses derived through the full dynamical analysis are in good agreement with the mean SPT based masses $M_{200}^{\rm SZ+\sigma}$ listed in Table 3.2. Regarding the precision of the constraints, it is clear that a composite cluster with ∼600 cluster members allows one to determine the dynamical mass with what is effectively a ∼ 15% uncertainty (∼8% for the full dataset, using ∼3000 tracers). In contrast, the NFW scale radius and the corresponding concentration are only weakly constrained.

Goodness of fit of dynamical models

We have also examined whether the best fit models are an adequate description of the data. To do this, we have created 1000 mock galaxy samples of similar size to the full galaxy sample by sampling the likelihood distribution in line of sight velocity and projected distance produced using the best fit model to the full galaxy sample. The best fit model has an MŁ anisotropy profile, and the Bayes factors of the other four models for the full galaxy sample are similar to those seen in Table 3.4 for the redshift subsamples of the data, indicating that all five anisotropy models fit similarly well. We then analyze each of these 1000 mock samples in the same way as the real data, examining the mean log(likelihood) per galaxy for the real data and the mocks. The likelihood of the real data is somewhat lower than the typical likelihood of the mocks. However, 6.1% of the mock samples have even lower likelihood than the real data. Thus, the dynamical models we are fitting are indeed an adequate description of the data.

3.4.2 Velocity dispersion anisotropy profiles

In the following we first present the results of our examination of mass and redshift trends (Section 3.4.2) and then discuss implications for a particular mass accretion model (Section 3.4.2) and then finally compare with previous studies of the velocity dispersion anisotropy profile (Section 3.4.2).

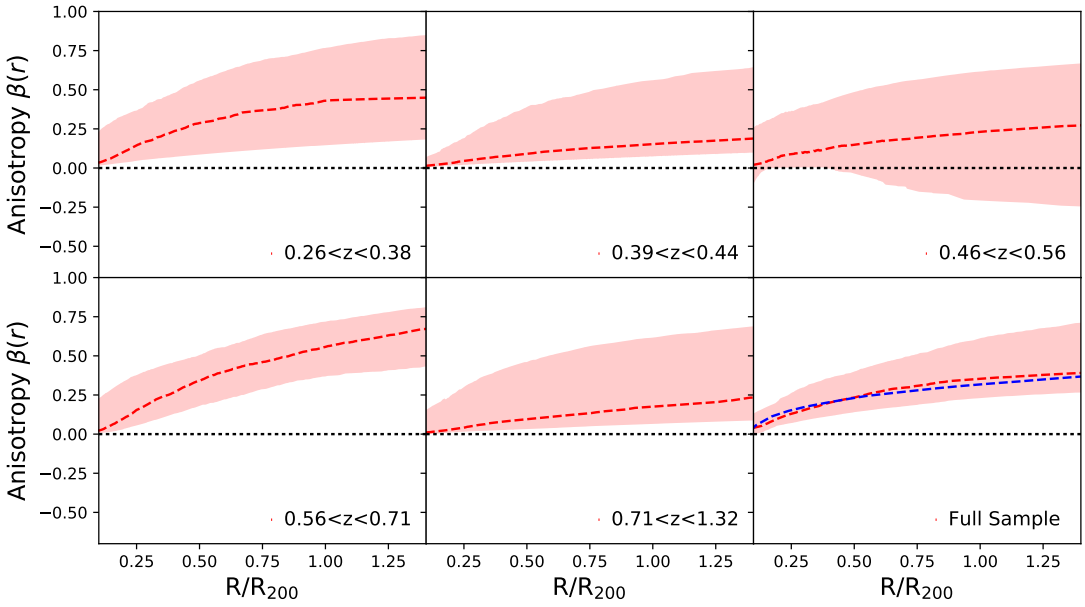

Figure 3.4: Velocity anisotropy profile $\beta(r)$ for each redshift bin. The red dashed line represents the profile obtained implementing a Bayesian model averaging, with the pink shaded region indicating the 1σ confidence region around this solution. There is no clear evidence of a redshift trend. The blue line in the lower-right panel (full sample) shows the result obtained when adopting the best fit NFW model and the Jeans equation inversion to solve for the anisotropy profile. This result is in good agreement with the model averaging result. Our analysis shows that passive galaxies preferentially move on nearly isotropic orbits close to the cluster center, and on increasingly radial orbits as one moves to the virial radius.

Constraints on Redshift and Mass Trends

Fig. 3.4 contains the measured anisotropy $\beta(r)$ profiles and their 1σ confidence regions for the five composite clusters in different redshift ranges together with the results from the full sample (lower, right-most panel). These profiles are obtained by using the posterior distribution in the anisotropy parameter θ_β extracted from each of the five anisotropy models. Specifically, for each of the models a large number of θ_β values are drawn, consistent with the posterior. Each value corresponds to an anisotropy profile. The number of θ_β values drawn for each model is weighted according to the Bayes factor. The sum of all these anisotropy profiles provides a measure of the probability distribution in the anisotropy profile value at each radius. The red line represents the median value of this distribution, while the shaded region is defined by the 16th and 84th percentiles of the distribution (1σ confidence region). As noted in the previous section, the mass profile model in all cases is an NFW with concentration and mass free to vary.

Our analysis indicates that the orbits of passive, red galaxies are nearly isotropic close to the cluster center, and become increasingly radial going towards larger radii, reaching a radial anisotropy $\beta \simeq 0.15 - 0.6$ at $R/r_{200} \simeq 1$. There is no clear evidence for a redshift trend in the anisotropy profile of the passive galaxy population out to $z \approx 1$. We have

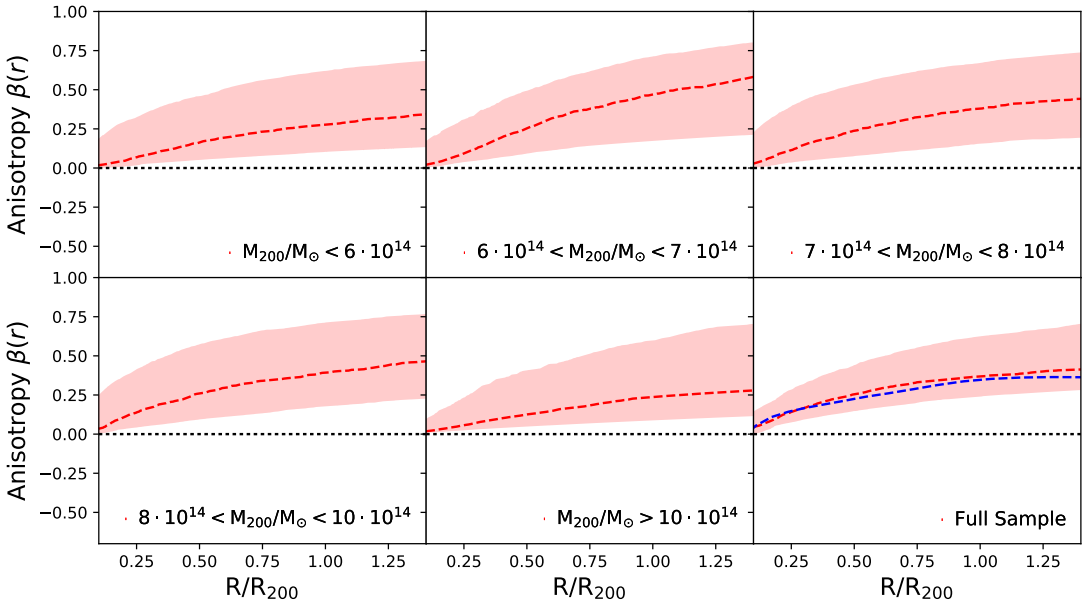

Figure 3.5: Velocity anisotropy profile $\beta(r)$ for each mass bin. As in Fig. 4, the anisotropy-redshiftred dashed line represents the profile obtained implementing a Bayesian model averaging, with the pink shaded region indicating the 1σ confidence region around this solution. The blue line in the lower-right panel (full sample) shows the result obtained when adopting the best fit NFW model and the Jeans equation inversion to solve for the anisotropy profile. There is no clear evidence of a mass trend.

carried out a similar analysis of five composite clusters built from the same cluster sample divided into mass bins rather than redshift bins, and we find no evidence for trends with mass, either. We show this result in Fig. 3.5.

For this reason we analyze also the full sample, providing our best available constraints. The orbital anisotropy varies from values consistent with zero in the cluster core to a value 0.4 ± 0.15 at the virial radius. For reference, anisotropy values of 0.4 correspond to tangential components of the velocity dispersion ellipsoid having amplitudes that are only 60 percent as large as the radial component. For the full sample we show (blue dashed line) also the anisotropy profile recovered using the best fit NFW parameters and using the Jeans equation to solve for the velocity dispersion anisotropy profile (Binney & Mamon, 1982; Solanes & Salvador-Sole, 1990; Biviano et al., 2013). This result is in good agreement with the solution recovered using the Bayesian model averaging over the five adopted anisotropy profiles.

Comparison with Two-Phase Accretion Model

The behavior of the anisotropy profile is consistent with the theoretical model discussed in Lapi & Cavaliere (2009), according to which the growth of structure proceeds in two phases: an early, fast accretion phase during which the cluster undergoes major merging events, and a second slower accretion phase involving minor mergers and smooth accretion (see, e.g.,

White, 1986; Zhao et al., 2003; Diemand et al., 2007). Lynden-Bell (1967) discusses how a dynamical system rapidly relaxes from a chaotic initial state to a quasi-equilibrium. The first stage of fast accretion provokes rapid changes in the cluster gravitational potential, inducing the collisionless components of the cluster to undergo violent relaxation, resulting in orbits that are more isotropic. Galaxies accreted by the cluster during the fast accretion phase would then be expected to exhibit approximately isotropic orbits, while galaxies accreted during the second phase would maintain their preferentially radial orbitals over longer timescales. Given that as the cluster accretes its mass and virial radius also grow, a two phase scenario like this would tend to lead to anisotropy profiles that are isotropic in the core and become more radial at larger radius.

However, one might also expect to see a time or redshift variation of the anisotropy profile, with typical galaxy orbits in high redshift clusters showing less of a tendency for radial orbits near the virial radius. The fact that our analysis shows no strong redshift trend in the anisotropy profiles is an indication that, in massive galaxy clusters, the passive galaxy population orbits are not changing significantly with cosmic time since $z \approx 1$. This suggests that the merging and relaxation processes responsible for the anisotropy profiles are underway at all cosmic epochs probed here. Other indicators of cluster merging have shown similar results; namely, the fraction of systems with disturbed X-ray morphologies (typically measured using centroid variations or ellipticities; Mohr et al., 1993) does not change significantly with redshift in samples of homogeneously SZE selected cluster samples (Nurgaliev et al., 2017; McDonald et al., 2017). This may suggest that the fast accretion phase of such a two phase model could be very short and happening primarily at redshifts above those probed by our sample. Clearly, a detailed examination of cosmological N-body simulations with sufficient volume to contain rare, massive clusters and sufficient mass resolution to ensure the survival of galaxy scale subhalos after accretion into the cluster is warranted. Moreover, a dynamical analysis like the one we have carried out that focuses on systems at redshift $z > 1$ would enable a more sensitive probe for time variation in the growth of structure.

Comparison with Previous Results

Previous studies of the velocity dispersion anisotropy conducted on passive cluster members at intermediate to high redshifts present hints of an anisotropy profile that is nearly isotropic close to the cluster center, and increasingly radial at larger radii. Biviano & Poggianti (2009), stacking 19 clusters between redshift $\approx 0.4 - 0.8$ with a mean mass of $\approx 3 \times 10^{14} M_{\odot}$, suggest radially anisotropic orbits. Biviano et al. (2016) reach the same conclusion when analysing a stacked sample of 10 clusters at $0.87 < z < 1.34$. Studying a single cluster, Annunziatella et al. (2016) show that the same trend is found for galaxies characterized by a stellar mass $M_{\star} > 10^{10} M_{\odot}$, while lower-mass galaxies move on more tangential orbits, avoiding small pericenters, presumably because those that cross the cluster center are more likelly to have been tidally destroyed. These results are consistent with ours. Our result, obtained through the analysis of a large sample of passive galaxies within a homogeneously selected sample of massive clusters over a wide redshift range and with

low scatter mass estimates, allows us to cleanly probe for redshift and mass trends in the velocity dispersion anisotropy profile.

Some published analyses carried out at lower redshifts ($z \lesssim 0.1$) than our sample show similar results. Wojtak & Łokas (2010) analyzed a sample of 41 nearby relaxed clusters, finding that galaxy orbits are isotropic at the cluster centers and more radial at the cluster virial radius. A similar result is obtained by Lemze et al. (2009) and Aguerri et al. (2017). However, other analyses show that the orbits of passive galaxies at these redshifts are more isotropic at all radii (Biviano & Katgert, 2004; Katgert et al., 2004a; Biviano & Poggianti, 2009; Munari et al., 2014), hinting at a possible change in galaxy orbits over time due to processes such as violent relaxation, dynamical friction, and radial orbital instability (Bellovary et al., 2008). At present, results from numerical simulations predict a range of behavior (Wetzel, 2011; Iannuzzi & Dolag, 2012; Munari et al., 2013), so further study is definitely needed. Extending our own observational analysis towards lower redshifts could also help clarify this picture.

3.4.3 Pseudo phase-space density profiles

The determination of the anisotropy profile $\beta(r)$ allows us to investigate the behavior of the Pseudo Phase-Space Density profile $Q(r)$ introduced in Section 3.1. According to numerical simulations of virialized halos (Taylor & Navarro, 2001; Dehnen & McLaughlin, 2005; Lapi & Cavaliere, 2009), there is a scaling between the density ρ and the velocity dispersion, best appreciated by considering the quantity $Q(r) = \rho/\sigma^3$. The profile of this quantity is found to follow a universal power-law of fixed slope, $\propto r^{-1.875}$. Remarkably, this is the same power-law predicted by the similarity solution of Bertschinger (1985) for secondary infall and accretion onto an initially overdense perturbation in an Einstein-de Sitter universe. In that work the authors found that the relaxation process is self-similar, meaning that each new shell falling in is virialized and adds a constant contribution to the resulting power-law density profile.

Because the velocity dispersion profile is constrained by the galaxies, we derive the PPSD profile using the number density profile of galaxies $\rho_{\rm gal}$ instead of the mass density profile. For each stacked cluster, we fix the Halo Occupation Number N_{200} of red galaxies to that found by H17 to find the central density ρ_0 of the NFW $\rho_{\rm gal}$ profile, such that

$$\rho_0 = \frac{N_{200}}{4\pi \int_0^{R_{200}} \frac{r^2}{r/r_s(1+r/r_s)^2} dr}. \tag{3.12}$$

We present the results obtained investigating both the total PPSD profile $Q_{\rm gal}(r) \equiv \rho_{\rm gal}/\sigma^3$ and the radial PPSD profile $Q_{\rm gal,r}(r) \equiv \rho_{\rm gal}/\sigma_r^3$, where $\rho_{\rm gal}$ is obtained as described above and $\sigma_r(r)$ is recovered using the following equation (van der Marel, 1994; Mamon & Łokas, 2005; Mamon et al., 2013)

$$\sigma_r^2(r) = \frac{1}{\nu(r)} \int_r^\infty \exp\left[2\int_r^s \beta(t)\frac{dt}{t}\right] \nu(s)\frac{GM(s)}{s^2} ds, \tag{3.13}$$

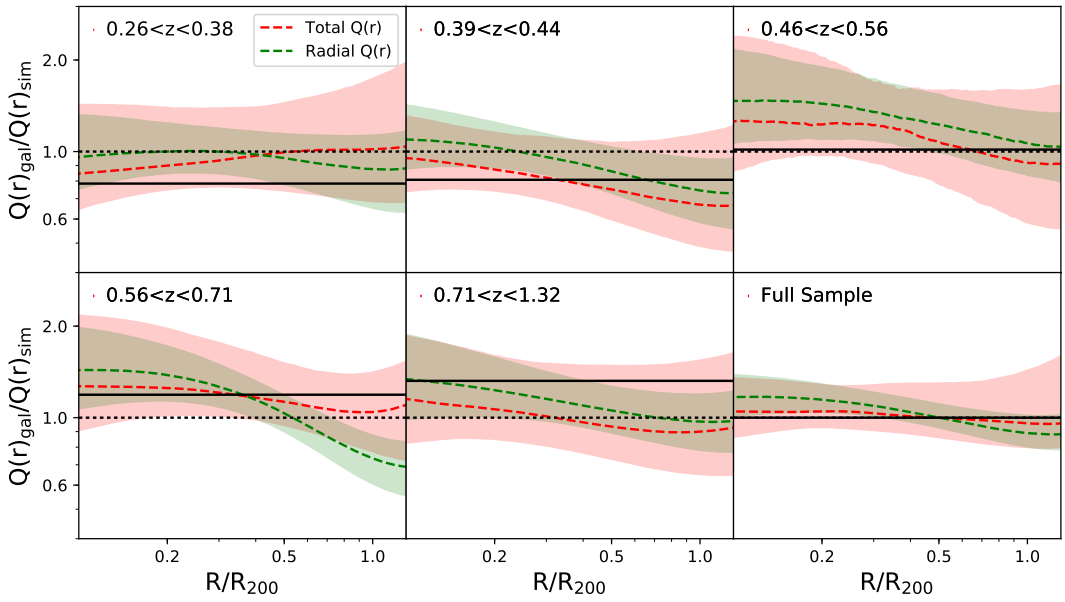

Figure 3.6: The ratios of the passive galaxy pseudo-phase-space density (PPSD) profiles $Q_{\rm gal}(r)$ to that of dark matter particles in simulations $Q_{\rm sim}(r)$ (Dehnen & McLaughlin, 2005) are shown in green for the total PPSD $\rho_{\rm gal}/\sigma^3$ and in red for the radial PPSD $\rho_{\rm gal}/\sigma_{\rm r}^3$ for five redshift bins and the full sample. Shading corresponds to the 1σ confidence region. The normalization of each $Q(r)$ profile (both of the galaxies and of the simulations) is fixed to that of the $Q_{\rm gal}(r)$ of the full sample, which is determined by fitting to the data. Thus, the deviations from $Q_{\rm gal}(r)/Q_{\rm sim}(r) = 1$ of the full sample are only driven by the deviations of the $Q_{\rm gal}(r)$ slope from the $Q_{\rm sim}(r)$ one, but in the other samples they are also driven by real normalization differences. The solid black line marks the expected amplitude in each redshift bin under the assumption of cluster self-similar redshift evolution. Both passive galaxy PPSD profiles are in reasonably good agreement with N-body simulations, and the data are roughly consistent with self-similar scaling with mass and redshift.

evaluated over an adequate grid of r, and $\sigma(r)$ is given by

$$\sigma(r) = \sqrt{3 - 2\,\beta(r)}\;\sigma_r(r). \tag{3.14}$$

Fig. 3.6 shows the ratio between the derived galaxy PPSD profiles and the fixed-slope best-fit relations $Q(r) \propto r^{-1.84}$ and $Q_r(r) \propto r^{-1.92}$, where the slopes are those obtained by Dehnen & McLaughlin (2005) by studying DM particles in numerically simulated halos, while the normalizations are fitted to the data of the full sample. A similar value for the slope of $Q(r)$ is found by Faltenbacher et al. (2007), a work based on numerical simulations and focused on the gas and DM entropy profiles in galaxy clusters. The similarity between radial profiles of the passive galaxy PPSD and simulated PPSD profiles is an indication that the passive galaxy radial distribution and kinematics are similar to those of the dark matter particles in those simulations.

We emphasize here that it is not possible for us to predict the PPSD profile for the underlying dark matter in our sample, because in general the blue or EL galaxy population

exhibits different kinematical properties than those of the passive, red or nEL galaxies. One might expect the full dark matter kinematics to exhibit a mix of the properties of the different galaxy populations. We note that because the measured concentration of the passive population (H17) is similar to the derived concentration of the mass profiles from our analysis, that if we replace the galaxy density profile with our derived mass density profile, the PPSD profiles have a very similar radial behavior.

In Fig. 3.6 we also show (solid black lines) the expected amplitude if there is self-similar behaviour of passive galaxy profiles and kinematics. That is, we have $Q(r) \equiv \rho/\sigma^3 \propto E(z)^2/[M(r)E(z)] = E(z)/M(r)$, where $H(z) = H_0 E(z)$ is the Hubble parameter. The passive galaxy PPSD amplitudes are in reasonably good agreement with the amplitudes expected under self-similarity, providing some additional evidence that this population is approximately self-similar.

In the previous sub-section, we mentioned that the anisotropy profiles could be the result of violent relaxation. This process, driven by gravity alone, would also tend to create a scale-invariant phase-space density. Relaxation into dynamical equilibrium would then lead the slope of the PPSD profile to approach a critical value, resulting in the particular form of the density profile for DM particles within simulated halos. An anisotropy profile isotropic in the inner regions and increasingly radial at larger radii gives the pseudo phase space density profile slope observed in numerical simulations. The agreement then suggests that the passive galaxies we analyze here have reached a similar level of dynamical equilibrium as the dark matter particles in the simulations, and that this is true for all redshifts up to $z \simeq 1$.

It is difficult to compare our results directly to previous studies, because most of those studies have tended to focus on $Q(r)$ and $Q_{\rm r}(r)$ using the inferred total matter density from their dynamical analyses rather than the galaxy density profile. Such an approach produces a mixed PPSD profile that contains both total matter and galaxy properties, because the dispersion profiles used are necessarily those of the galaxies rather than the dark matter. This complicates the interpretation, because a mismatch between the observed and simulated PPSD profiles could be because the total matter density profile doesn't match the simulated dark matter profile or it could mean that the velocity dispersion profile of the dark matter and a particular galaxy population do not match. In general, given that different galaxy populations tend to exhibit different kinematic properties (i.e. radial distributions and velocity dispersion and anisotropy profiles), one would not expect the mixed PPSD profiles defined in this way to agree for the different populations.

Indeed, Biviano et al. (2013) measured $Q(r)$ using the anisotropy profile of the galaxies and the NFW mass density profile inferred from their analysis, separately for star forming (SF) and passive galaxies in a single cluster. They found that their observed PPSD profile agreed with simulations only for the passive galaxy population. They argue that passive members have undergone violent relaxation and have reached dynamical equilibrium, while SF members have not yet reached equilibrium. But another possible interpretation would be that the Jeans analyses of both populations are equally accurate, but that the passive population exhibits a velocity dispersion and anisotropy profile more similar to that of the simulated dark matter. A more recent study came to similar conclusions (Munari et al.,

2015).

On the other hand, a recent study of the nearby cluster Abell 85 (Aguerri et al., 2017) shows that $Q_r(r)$ follows the theoretical power-law form independent of the galaxy colour or luminosity, concluding that all the different families of galaxies under study reached a virialized state. Given the discussion above, this agreement in the mixed PPSD profile derived from different galaxy populations also indicates that the different populations must have similar kinematics (i.e. velocity dispersion and anisotropy profiles). In their study, they emphasize that the anisotropy profiles of the blue and red galaxies are different, which would make the agreement in $Q_{\rm r}(r)$ surprising. However, Fig. 3 in their paper suggests that the anisotropy profiles have similar character (isotropic in the center, more radial at larger radius) and present evidence for inconsistency that is weak ($\leq 2\sigma$).

Bin	Redshift range	$M_{200}^{\rm dyn} \mid M_{200}^{\rm SZ+\sigma}$ $[10^{14} M_\odot]$	$M_{200}^{\rm dyn} \mid M_{200}^{\rm SZ+Planck}$ $[10^{14} M_\odot]$	$\frac{M_{200}^{\rm dyn} \mid M_{200}^{\rm SZ+\sigma}}{\langle M_{200}^{\rm SZ+\sigma} \rangle}$	$\frac{M_{200}^{\rm dyn} \mid M_{200}^{\rm SZ+Planck}}{\langle M_{200}^{\rm SZ+Planck} \rangle}$	η	η'
1	0.26-0.38	$9.44^{+1.70}_{-1.65}$	$9.29^{+1.96}_{-1.45}$	$1.02^{+0.18}_{-0.18}$	$0.86^{+0.15}_{-0.15}$	$0.77^{+0.28}_{-0.13}$	$1.21^{+0.17}_{-0.36}$
2	0.39-0.44	$10.57^{+1.93}_{-1.55}$	$10.41^{+1.88}_{-1.19}$	$1.11^{+0.20}_{-0.16}$	$0.91^{+0.17}_{-0.13}$	$0.92^{+0.17}_{-0.20}$	$1.24^{+0.21}_{-0.27}$
3	0.46-0.56	$7.42^{+1.58}_{-0.92}$	$7.14^{+1.49}_{-0.88}$	$0.93^{+0.20}_{-0.11}$	$0.74^{+0.16}_{-0.09}$	$0.74^{+0.15}_{-0.09}$	$0.97^{+0.23}_{-0.12}$
4	0.56-0.71	$7.31^{+1.13}_{-0.62}$	$7.36^{+1.01}_{-0.78}$	$1.00^{+0.15}_{-0.08}$	$0.73^{+0.11}_{-0.06}$	$0.79^{+0.10}_{-0.10}$	$1.10^{+0.13}_{-0.13}$
5	0.71-1.32	$6.20^{+0.85}_{-0.88}$	$5.95^{+0.83}_{-0.75}$	$0.82^{+0.11}_{-0.12}$	$0.55^{+0.07}_{-0.08}$	$0.63^{+0.13}_{-0.08}$	$0.90^{+0.18}_{-0.11}$
-	0.26-1.32	$8.71^{+0.52}_{-0.80}$	$8.50^{+0.59}_{-0.67}$	$1.05^{+0.06}_{-0.10}$	$0.81^{+0.05}_{-0.07}$	$0.81^{+0.06}_{-0.06}$	$1.14^{+0.06}_{-0.07}$

Table 3.6: Comparisons of dynamical masses from composite clusters calculated using different initial masses. From left to right the columns contain the redshift range of the cluster sample, the derived dynamical mass given an initial SPT plus velocity dispersion mass, the derived dynamical mass given an SPT + Planck initial mass, and the ratios of these dynamical to initial masses in each case. Finally, we report the constraints on η and η' as described in Section 3.4.4.

3.4.4 Dynamical mass constraints

We use the dynamical information to study the masses of these clusters in two different ways. In Section 3.4.4, we analyze the composite clusters and the consistency of the dynamical masses when using different initial masses to scale the galaxy observables. In Section 3.4.4, we examine the differences between the dynamical masses and the SPT+Planck masses $M_{200}^{\rm SZ+Planck}$ using a different approach where the scaling values for each cluster r_{200} and v_{200} are altered self-consistently in each iteration of the Markov chain to reflect the SPT+Planck masses scaled by a free parameter η, defined in equation (3.15).

Mass constraints on the composite clusters

Constructing the composite clusters (described fully in Section 3.3.3) requires a scaling of our galaxy observables, $v_{\rm rf}$ and R, by estimates of the virial radius and velocity r_{200} and v_{200}

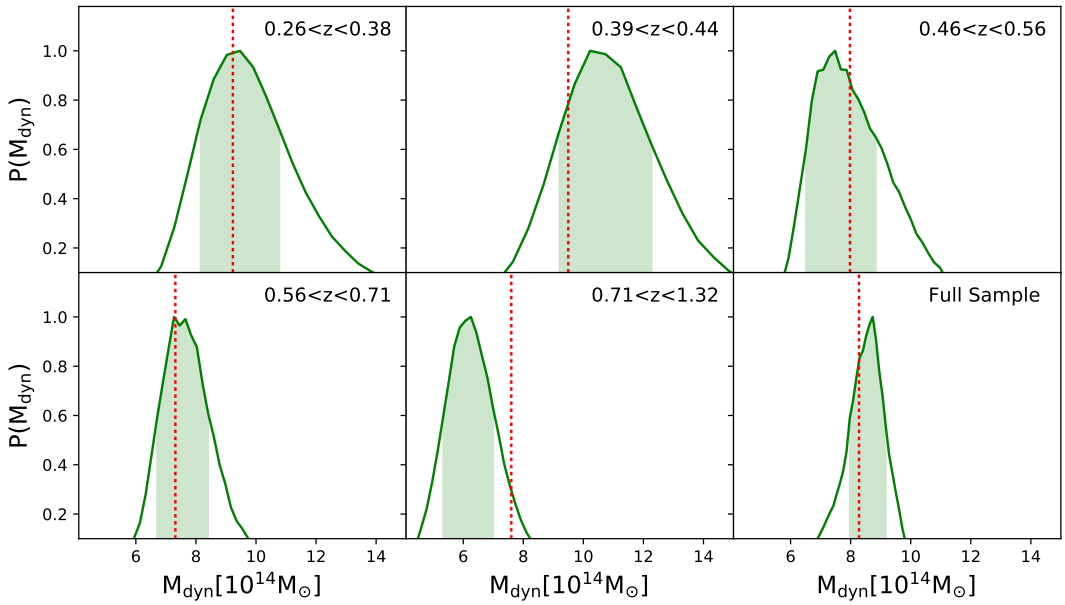

Figure 3.7: Marginalized distribution of the dynamical masses. Each panel corresponds to a different redshift range, and the final panel shows the results of the analysis of the full sample. In green we highlight the 1σ confidence region. The red line represents the mean SPT+σ mass $\langle M_{200}^{\mathrm{SZ}+\sigma}\rangle$ for the clusters in the bin, weighted by the number of member galaxies in the individual clusters. There is a good agreement between the dynamical masses and the originally inferred SPT masses in all cases (see Table 3.6).

respectively, and therefore requires an initial mass estimate. As described above, this is a potential problem, because there is currently a ∼25 percent shift between the SPT cluster masses derived using cluster counts and velocity dispersion information, and the masses derived using the cluster counts and external cosmological priors from Planck (Bocquet et al., 2015; de Haan et al., 2016). Thus, we examine the dynamical mass constraints in each case: (1) those derived using initial masses derived from the cluster counts and velocity dispersion measurements $P(M_{200}^{\mathrm{dyn}}\,|M_{200}^{\mathrm{SZ}+\sigma})$ and (2) those derived using initial masses from the cluster counts and external cosmological constraints from Planck $P(M_{200}^{\mathrm{dyn}}\,|M_{200}^{\mathrm{SZ+Planck}})$.

In Fig. 3.7 we display the marginalized distribution of the dynamical masses obtained with our analysis where the $M_{200}^{\mathrm{SZ}+\sigma}$ masses were adopted for the initial scaling. The green regions mark the 1σ confidence regions, and the red lines represent the mean initial masses derived from the cluster counts and velocity dispersion measurements $\langle M_{200}^{\mathrm{SZ}+\sigma}\rangle$, where the masses from each cluster are weighted by the number of galaxy velocities available for that cluster. There is good agreement between the dynamical masses and the initial masses in all redshift bins and also for the full sample (lower, right-most panel). The second column of Table 3.6 contains the measurement results and uncertainties for each subset. Characteristic dynamical mass uncertainties are at the ∼15 percent level for individual subsamples and at the ∼8 percent level for the full sample. These are quite encouraging mass constraints, given that they are marginalized over the velocity dispersion anisotropy profile uncertainties.

We find that by fixing the concentration and radial anisotropy parameters to their best fit values when fitting for mass in a composite cluster with 600 tracers, the resulting mass uncertainty is not significantly impacted. This suggests that our uncertainties in the individual redshift bins are not dominated by the freedom in mass and anisotropy profiles. A similar test on the composite cluster built from the full dynamical sample leads to a $\sim$5% mass uncertainty, which is comparable to what we find when using a single anisotropy model, before performing the Bayesian model averaging. This is an interesting result when taken together with the discussion of velocity dispersion based mass estimates in Sifón et al. (2016), where the scaling presented suggests that with samples of 600 dynamical tracers mass estimates should be closer to $\sim$7% accurate rather than the 15% we recover. Further examination of the assumptions built into the dynamical mass measurements using velocity dispersions and full Jean analysis modeling is warranted and is planned for a future analysis.

As discussed in Section 3.2.1, current estimates from studies of clusters in numerical simulations indicate there are remaining systematic uncertainties associated with MAMPOSSt analysis at the 10% level.

To test the stability of the recovered dynamical masses to the initial input masses used for scaling, we perform the same analysis using the SPT+Planck cluster masses. The third column of Table 3.6 shows these results. These dynamical masses with the different initial masses are quite close to the values derived with the other set of initial masses. This shows that there is no strong dependence of the dynamical mass on the initial mass. This is because any change in the masses used for rescaling the cluster observables during stacking will impact, on average, the individual cluster masses and the final mean mass in the bin in a similar manner. The overall scale of the dynamical data in projected radius and LOS velocity remains approximately invariant.

As columns four and five of Table 3.6 make clear, the dynamical masses, while being in good agreement with the cluster counts plus velocity dispersion masses $M_{200}^{\mathrm{SZ}+\sigma}$, exhibit some discordance with the cluster counts plus external cosmological constraint masses $M_{200}^{\mathrm{SZ+Planck}}$. While the three lowest redshift bins show no significant disagreement, the upper two redshift bins show masses that are only 73 percent and 55 percent as large as the SPT+Planck masses (offsets that are statistically significant at the 2.5σ and 6.5σ levels, respectively). The full sample has a dynamical mass that is only 80 percent of the SPT+Planck masses, a difference that is significant at the 3.8σ level (statistical only). The direction and scale of this mass shift is similar to that highlighted already in Bocquet et al. (2015). However, with our analysis we are able to show that this discrepancy seems to grow with redshift.

Comparison with SZE based masses

To examine this discrepancy more carefully, we use the dynamical analysis to examine the masses of these clusters and compare them to the masses derived separately from the SPT cluster counts in combination with external cosmological constraints from the Planck CMB anisotropy. Rather than using the composite clusters, we analyze individual clusters,

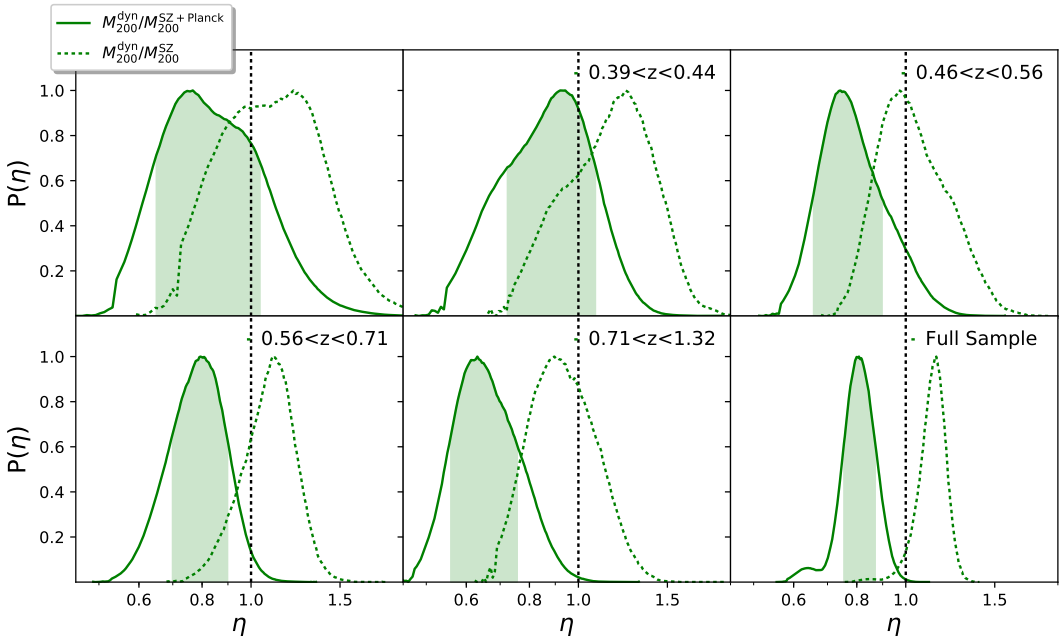

Figure 3.8: Posterior distribution of $\eta = M_{200}^{\rm dyn}/M_{200}^{\rm SZ+Planck}$ (solid green lines) arising from a dynamical analysis of each cluster subsample. The green region shows the 1σ region, while the vertical dotted black line marks the value $\eta = 1$. These distributions show that in the two high redshift bins and for the full sample there is disagreement between the dynamical masses and the SPT+Planck calibrated cluster masses. For the full sample the discrepancy is 1.9σ significant when including estimates of the systematic uncertainties. In the highest redshift bin, the discrepancy is 2.6σ. The dashed green lines show the estimated posterior for $\eta' = M_{200}^{\rm dyn}/M_{200}^{\rm SZ}$, where $M_{200}^{\rm SZ}$ represents the de Haan et al. (2016) masses calibrated using the SPT mass function and Y_X measurements for many of the systems. In contrast to the SPT+Planck masses, these masses are in good agreement (1σ offset) with the dynamical masses.

combining the likelihoods from each cluster and exploring constraints on an overall mass scaling parameter η, that is defined as

$$\eta = \frac{M_{200}^{\rm dyn}}{M_{200}^{\rm SZ+Planck}}. \tag{3.15}$$

We do this by running MAMPOSSt for each individual cluster in our sample. We calculate the posterior distribution of η by using a multimodal nested sampling algorithm, namely MultiNest (Feroz & Hobson, 2008; Feroz et al., 2009, 2013), which provides us with the evidence for each model, and allows us to perform a Bayesian model averaging over different subsets of clusters.

Fig. 3.8 contains a plot of the posterior distributions of η from our analysis within each redshift bin and for the full sample. Results are largely consistent with the results from the composite clusters. Column six of Table 3.6 contains the best fit η values and associated uncertainties. The preferred value for the full sample is $\eta = 0.81 \pm 0.06 \pm 0.08$. The constraint is followed by a statistical uncertainty and then a systematic uncertainty.

As already discussed in Section 3.2.1, studies of dynamical tracers drawn from clusters in structure formation simulations indicate that MAMPOSSt derived dynamical masses has systematic uncertainties of $\approx 10\%$ (see Mamon et al., 2013). This number comes from an analysis of tracers lying within a sphere of radius r_{100} that are then used to estimate the virial radius r_{200}. For the systematic uncertainty presented above, we have therefore adopted as a Gaussian with $\sigma = 10\%$ centered at no bias.

If one combines the statistical and systematic uncertainty in quadrature, the implication would be a difference at the 1.9σ level. As mentioned already, this tendency for the dynamical masses to be lower than those masses derived from the cluster counts in combination with external cosmological constraints is consistent with the tendencies seen previously (Bocquet et al., 2015) using simply dispersions and the cluster mass function (see also Rines et al., 2016; Sifón et al., 2016). More recent weak lensing analyses also support the lower mass scale of SPT clusters (Dietrich et al., 2017; Stern et al., 2018).

To emphasize, Fig. 3.8 also shows (dotted line) the distribution of

$$\eta' = \frac{M_{200}^{\rm dyn}}{M_{200}^{\rm SZ}}, \tag{3.16}$$

where $M_{200}^{\rm SZ}$ are masses calibrated from a cosmological analysis carried out in de Haan et al. (2016) using the X-ray mass proxy Y_x and the abundance of clusters as a function of redshift, Y_x and the SZE mass proxy ξ, without the inclusion of the external Planck cosmological constraints. These results indicate that the dynamical masses are in good agreement with the $M_{200}^{\rm SZ}$ masses at all redshifts and for the full sample. The η' distribution for the full sample prefers a value of $\eta' = 1.14 \pm 0.07 \pm 0.11$, indicating no disagreement.

To summarize, our dynamical mass measurements, which are derived using only dynamical information and no information from the mass function or cluster counts, are in good agreement with masses derived using information from the the cluster counts together with additional information from either velocity dispersions or from X-ray Y_X measurements that have been externally calibrated. However, our mass measurements exhibit moderate disagreement with those masses obtained similarly but when also adopting external cosmological priors from Planck CMB anisotropy. Progress in testing these two mass scales would require better control of the systematic uncertainties in the dynamical masses (Mamon et al., 2013). The agreement between the dynamical and the Planck based masses is best at low redshift, with the dynamical masses preferring ever smaller η with increasing redshift. In the highest redshift bin ($0.71 \leq z \leq 1.32$) we measure $\eta = 0.63^{+0.13}_{-0.08} \pm 0.06$, discrepancy at the 3σ level.

3.4.5 Impact of disturbed clusters

One assumption in applying the Jeans equation to analyze our sample is that the galaxies we are analyzing are in approximate dynamical equilibrium. Thus, it would seem important to remove the obviously disturbed clusters— those undergoing or having undergone recent, major mergers. As discussed in Mohr et al. (1993), the asymmetry and isophotal ellipticity

Subsample	$\langle z \rangle$	c_{200}	$M_{200}^{\rm dyn}\,\vert M_{200}^{\rm SZ+\sigma}$ $[10^{14} M_\odot]$	$\frac{M_{200}^{\rm dyn}\,\vert M_{200}^{\rm SZ+\sigma}}{\langle M_{200}^{\rm SZ+\sigma} \rangle}$
relaxed	0.57	$6.4^{+1.4}_{-2.4}$	$7.0^{+0.8}_{-0.9}$	$0.8^{+0.1}_{-0.1}$
un-relaxed	0.6	$4.5^{+1.2}_{-1.1}$	$9.5^{+1.0}_{-1.2}$	$1.0^{+0.1}_{-0.1}$

Table 3.7: Sensitivity of dynamical mass measurements to the dynamical state of clusters. We compare the masses for those clusters exhibiting large X-ray surface brightness asymmetries (un-relaxed) and those exhibiting small asymmetries (relaxed). Columns list the subsample, the mean redshift, concentration, dynamical mass given initial scaling using SPT dispersion based masses, and ratio of the dynamical mass to the SZE dispersion based mean mass for the subsample.

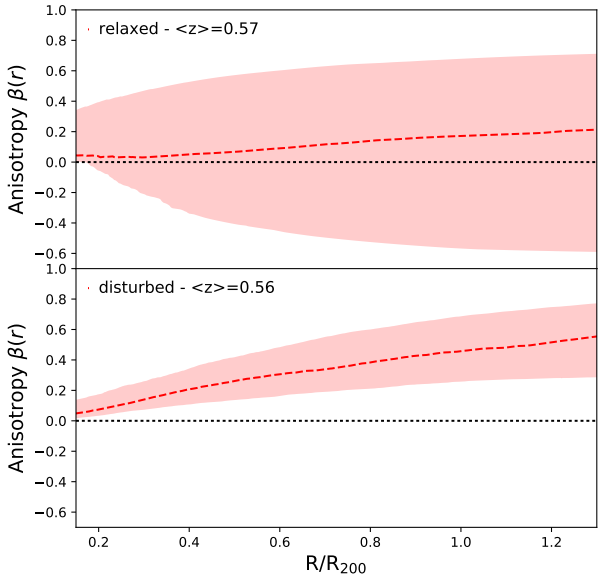

Figure 3.9: Velocity anisotropy profile $\beta(r)$ recovered when separately analyzing relaxed and un-relaxed clusters. For the latter, the anisotropy profile indicates that the galaxies in the clusters are moving on nearly isotropic orbits near the center and increasingly radial orbits out toward the virial radius R_{200}. The recovered anisotropy profile for the relaxed clusters is consistent with these results, but exhibits much larger uncertainties that allow for a much broader range of galaxy orbits.

of the X-ray surface brightness distribution provide information about the merger state of galaxy clusters, which is generally superior to the constraints possible from the galaxy distribution (Geller & Beers, 1982) due to higher signal to noise. However, neither of these measures are sensitive to mergers along the line of sight, where the galaxy velocity distribution typically provides more information (Dressler & Shectman, 1988).

Merger signatures from X-ray cluster surface brightness distributions were extracted

uniformly from a large X-ray flux limited sample of nearby clusters, indicating that over half of them exhibit statistically significant centroid variations and that the bulk of them are elliptical (Mohr et al., 1995). That study provided clear evidence that even massive clusters at low redshift are still undergoing continued accretion of subclusters. This observation is in agreement with expectations from structure formation within our standard ΛCDM model. We can use a similar approach to identify major mergers (not along the line of sight) in our cluster sample. Nurgaliev et al. (2017) quantified the X-ray morphology of a subsample of 90 SZE selected galaxy clusters using a measure of the photon asymmetry A_{phot} (Nurgaliev et al., 2013) closely related to the centroid variation (Mohr et al., 1993, 1995) and power ratios (Buote & Tsai, 1995, 1996).

To test the dependence of our results on the dynamical state of the clusters, we adopt these A_{phot} values as a measure of departures from equilibrium and separately analyze relaxed and disturbed clusters. We take here relaxed clusters to be those with $A_{phot} \leq 0.2$. Out of our 110 cluster sample, 68 (2152 spectra) have measured A_{phot} and 39 of these (1258 galaxies) are classified as relaxed. In Table 3.7 we show the results for the mass $M_{200}^{\rm dyn}$ and concentration c_{200} having performed the analysis on the stacked relaxed/disturbed populations. We find that the $M_{200}^{\rm SZ+\sigma}$ and the dynamical masses are in agreement at $\approx 1.5\sigma$ level for the relaxed sample and at the 1σ level for the disturbed sample (statistical only).

Fig. 3.9 shows the recovered anisotropy profiles for the two subsamples. The anisotropy profile recovered for the disturbed clusters indicates that the galaxies are moving on roughly isotropic orbits in the center and increasingly radial orbits at large distance from the cluster core. This is consistent with the behavior seen in the total sample and most subsamples. On the other hand, we find that the anisotropy profile of relaxed clusters, while consistent with this behavior, exhibits much larger uncertainties that allow also for very different behavior, including simple isotropic orbits at all radii. A comparison shows that in the case of the relaxed sample one of the anisotropy models, namely the $\beta_O(r)$ one, which has anisotropy of opposite sign in the center and at large radii, has the highest Bayes factor. In the case of this sample, the preference is for radial orbits near the center and somewhat tangential orbits at larger radius. This is responsible for the extension of the uncertainties to anisotropies of ~ -0.4 at $R/R_{200} \sim 1$. We note that this kind of behavior is also shown in the upper right panel of Fig. 3.4, which is the redshift subsample that also prefers the β_O model (see Table 3.4).

The differences in the character of these results is intriguing and deserves further study with larger dynamical samples in cluster ensembles that have associated measurements of substructure. We note, however, that the results presented in this section on the velocity anisotropy profiles, the PPSD profiles and the cluster halo masses have been compared to studies of clusters formed in numerical structure formation simulations (e.g. Navarro et al., 1996; Dehnen & McLaughlin, 2005; Faltenbacher et al., 2007; Mamon et al., 2013) in which cluster substructure is generic. Observationally, substructure has been established as generic to the real cluster population at low redshift for a timespan approaching four decades (Geller & Beers, 1982). With a uniform selection of cluster subsamples defined in

similar ways in simulations and the real world, it should be possible in future analyses to sharpen studies like ours to measure differences in galaxy orbital characteristics associated with X-ray substructure, presence of cool cores and so on.

3.5 Conclusions

We present a dynamical analysis of 110 SZE selected galaxy clusters from the SPT-SZ survey with redshifts between 0.2 to 1.3, that have an associated spectroscopic sample of more than 3000 passive galaxies. We examine subsets of this cluster sample in redshift and mass, each comprising ∼600 cluster members. These subsets are either combined to form composite clusters or are analyzed individually using a Jeans equation based code called MAMPOSSt (Mamon et al., 2013) that allows one to adopt different parametric models for the mass profile, galaxy profile and velocity dispersion anisotropy profile. In our analysis we adopt an NFW mass profile, and use the measured concentration of the red sequence galaxy population from a complete subsample of the SPT SZE selected cluster sample (H17), and employ five different velocity dispersion anisotropy profiles (see Section 3.2.2). We perform Bayesian model averaging to combine results from the different dispersion anisotropy models, because none of the five models are excluded by the data.

The velocity dispersion anisotropy profiles show the same radial features at all redshifts: orbits are isotropic near the center and increasingly radial at larger radii. We also find no variations with cluster mass. The radial variation is broadly consistent with that seen in a recent analysis of near-infrared selected clusters at $z \sim 1$ (Biviano et al., 2016) and also studies at low redshift ($z \lesssim 0.1$; Lemze et al., 2009; Wojtak & Łokas, 2010). These trends of anisotropy with radius resemble those of DM particles in halos extracted from cosmological numerical simulations (Mamon & Łokas, 2005; Mamon et al., 2010b, 2013, and references therein). The absence of a redshift trend is inconsistent with the results presented in Biviano & Poggianti (2009), where they report that passive galaxy orbits are becoming more isotropic over time. The absence of a redshift trend in the velocity anisotropy profiles suggests that the process of infall and relaxation for the passive galaxy population is occurring similarly at all redshifts since at least $z \sim 1$.

We measure the pseudo-phase-space density (PPSD) profiles $Q_{\rm gal}(r)$ and $Q_{\rm gal,r}(r)$, using quantities derived from cluster galaxies. We find good agreement with theoretical predictions from N-body simulations of DM particles (Dehnen & McLaughlin, 2005). We examine whether the amplitude of the profile scales as expected with redshift and mass under the assumption of self-similarity, finding that they do. To the extent that the PPSD profile provides constraints on the equilibrium nature of the galaxy dynamics, the good agreement with simulations suggests that galaxies behave approximately as collisionless particles and are as relaxed as the DM particles in halos forming within cosmological structure formation simulations. Moreover, the lack of evidence for redshift trends in the power law index of the PPSD profiles suggests again that the passive galaxy population in clusters is dynamically similar at all redshifts and mass ranges probed in our study.

We carry out a consistency check between our dynamical masses $M_{200}^{\rm dyn}$, which are

marginalized over uncertainties in the velocity dispersion anisotropy profiles, with masses calibrated using the SPT cluster counts with and without strong external cosmological priors. We find that our masses are smaller than those derived with strong external cosmological priors $M_{200}^{\mathrm{SZ+Planck}}$ by $\eta = 0.81 \pm 0.06 \pm 0.08$, corresponding to a 1.9σ discrepancy when systematic uncertainties in our dynamical masses are included. Moreover, our analysis shows that the agreement is best at low redshift, while lower values of η are preferred at higher redshift. In the highest redshift bin the best fit mass ratio is $\eta = 0.63^{+0.13}_{-0.08} \pm 0.06$, which corresponds to a disagreement at the 2.6σ level.

In addition, we find good agreement, $\eta' = 1.14 \pm 0.07 \pm 0.11$, between our dynamical masses and those masses extracted from the SPT cluster counts in combination with 82 externally calibrated X-ray Y_X mass estimates (de Haan et al., 2016) when the cosmological parameters are allowed to vary (and also those masses calibrated in combination with 63 velocity dispersions; see Bocquet et al., 2015). Our mass constraints are also consistent with those from related studies of SPT selected clusters, using both weak lensing magnification (Chiu et al., 2016a) and tangential shear (Dietrich et al., 2017; Schrabback et al., 2018; Stern et al., 2018).

Using *Chandra* X-ray data, we examine the impact of the dynamical state of the clusters on our dynamical analysis by separately analyzing relaxed and un-relaxed clusters. We find dynamical masses to be in good agreement with our combined sample for both the relaxed and un-relaxed clusters. Concerning the anisotropy profiles, we find that, for the disturbed sample, the shape of the orbits resembles the one seen in the total sample and most subsamples. On the other hand, the anisotropy profile of relaxed clusters, while still consistent with this behavior, exhibits much larger uncertainties that allow also for isotropic orbits at all radii. Further investigation with larger dynamical samples in cluster ensembles is required in order to understand the different behaviour of these objects.

As a next step, our analysis can be extended to cluster samples that include many low mass systems. One such sample that is being analyzed presently has been defined in the project known as SPIDERS (SPectroscopic IDentification of eROSITA Sources, Clerc et al., 2016), an optical spectroscopic survey of X-ray-selected galaxy clusters discovered in ROSAT and *XMM-Newton* imaging. Another sample is being built up through spectroscopic observations of optically selected clusters within the Dark Energy Survey. Longer term, we expect deep spectroscopic followup of SZE and X-ray selected clusters to provide ever larger galaxy samples that include both emission line and passive galaxies. These samples will allow cluster masses to be constrained in a redshift regime where weak lensing is challenging, while also enabling studies of the kinematic relationship between cluster emission line and passive galaxies out to redshifts well beyond 1.

Acknowledgments

We thank Cristóbal Sifón for providing useful feedback. We acknowledge the support by the DFG Cluster of Excellence "Origin and Structure of the Universe", the Transregio program TR33 "The Dark Universe" and the Ludwig-Maximilians University. The South Pole Telescope is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-1125897 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation and the Gordon and Betty Moore Foundation grant GBMF 947. The Melbourne group acknowledges support from the Australian Research Council's Discovery Projects funding scheme (DP150103208). DR is supported by a NASA Postdoctoral Program Senior Fellowship at NASA's Ames Research Center, administered by the Universities Space Research Association under contract with NASA. Work at Argonne National Laboratory was supported under U.S. Department of Energy contract DE-AC02-06CH11357. AB acknowledges the hospitality of the LMU, and partial financial support from PRIN-INAF 2014 "Glittering kaleidoscopes in the sky: the multifaceted nature and role of Galaxy Clusters?", P.I.: Mario Nonino. BB has been supported by the Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. AS is supported by the ERC-StG "ClustersXCosmo", grant agreement 716762.

Chapter 4

Mass Calibration of the CODEX Cluster Sample using SPIDERS Spectroscopy - I. The Richness-Mass Relation

R. Capasso[1,2], J. J. Mohr[1,2,3], A. Saro[1,2,4], A. Biviano[4], N. Clerc[3,5], A. Finoguenov[3,6], S. Grandis[1,2], C. Collins[7], G. Erfanianfar[3], S. Damsted[6], C. Kirkpatrick[6], A. Kukkola[6]

accepted for publication in Monthly Notices of the Royal Astronomical Society (MNRAS), ArXiv:1812.06094

[1]Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
[2]Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
[3]Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 85748 Garching, Germany
[4]INAF-Osservatorio Astronomico di Trieste via G.B. Tiepolo 11, 34143 Trieste, Italy
[5]IRAP, Université de Toulouse, CNRS, UPS, CNES, Toulouse, France
[6]Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2a, FI-00014 Helsinki, Finland
[7]Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool, L3 5RF, UK

ABSTRACT

We use galaxy dynamical information to calibrate the richness–mass scaling relation of a sample of 428 galaxy clusters that are members of the CODEX sample with redshifts up to $z \sim 0.7$. These clusters were X-ray selected using the ROSAT All-Sky Survey (RASS) and then cross-matched to associated systems in the redMaPPer catalog from the Sloan Digital Sky Survey. The spectroscopic sample we analyze was obtained in the SPIDERS program and contains ∼7800 red member galaxies. Adopting NFW mass and galaxy density profiles and a broad range of orbital anisotropy profiles, we use the Jeans equation to calculate halo masses. Modeling the scaling relation as $\lambda \propto \mathrm{A}_\lambda {M_{200\mathrm{c}}}^{\mathrm{B}_\lambda}(1+z)^{\gamma_\lambda}$, we find the parameter constraints $\mathrm{A}_\lambda = 38.6^{+3.1}_{-4.1} \pm 3.9$, $\mathrm{B}_\lambda = 0.99^{+0.06}_{-0.07} \pm 0.04$, and $\gamma_\lambda = -1.13^{+0.32}_{-0.34} \pm 0.49$, where we present systematic uncertainties as a second component. We find good agreement with previously published mass trends with the exception of those from stacked weak lensing analyses. We note that although the lensing analyses failed to account for the Eddington bias, this is not enough to explain the differences. We suggest that differences in the levels of contamination between pure redMaPPer and RASS+redMaPPer samples could well contribute to these differences. The redshift trend we measure is more negative than but statistically consistent with previous results. We suggest that our measured redshift trend reflects a change in the cluster galaxy red sequence fraction with redshift, noting that the trend we measure is consistent with but somewhat stronger than an independently measured redshift trend in the red sequence fraction. We also examine the impact of a plausible model of correlated scatter in X-ray luminosity and optical richness, showing it has negligible impact on our results.

4.1 Introduction

The formation and evolution of galaxy clusters is governed by the complex interplay between the gravity-induced dynamics of collapse and the baryonic processes associated with galaxy formation. Galaxy clusters, thus, constitute unique laboratories for both astrophysics and cosmology. On one side, the abundance of these objects as a function of mass and redshift is a well established cosmological probe (e.g., White et al., 1993a; Haiman et al., 2001; Mantz et al., 2015; de Haan et al., 2016). On the other side, the observation of the evolution of galaxy properties in clusters provide us with information on galaxy formation, their assembly history, and the correlation between their evolution and environment (e.g., Dressler, 1984; de Propris et al., 1999; Mei et al., 2009; Muzzin et al., 2012; Hennig et al., 2017; Strazzullo et al., 2018; Capasso et al., 2019).

Of primary importance to both types of studies are accurate mass estimates and large samples of clusters with well understood selection. For cosmological studies that adopt the halo mass function this is obvious, but for galaxy population studies it is equally important, because galaxy properties vary with clustercentric distance, and thus to compare properties of clusters across a range of mass and redshift, it is crucial to be able to adopt a meaningful overdensity radius such as $r_{200\mathrm{c}}$, which corresponds to the radius at which the mean enclosed density is 200 times the critical density and is thus trivially derived from the corresponding mass $M_{200\mathrm{c}}$. Adopting an overdensity radius reveals cluster regularity or approximate self-similarity in structure formation simulations (e.g, Navarro et al., 1997) and has also revealed regularity in studies of real clusters (e.g. Pratt et al., 2007).

A good understanding of the mass–observable relation that links the mass of galaxy

clusters to readily obtainable observables such as the optical richness λ is then more than a convenience. It enables both cosmological and structure formation studies on large cluster ensembles. Within this context, uncertainties on cluster masses include the measurement uncertainties on the observable, the intrinsic scatter in the observable at fixed mass and redshift and the uncertainties on the parameters of the mass–observable relation. The latter can be controlled through calibration.

Different mass constraints have been used to calibrate the mass–observable relation for cluster ensembles, each with its advantages and disadvantages. Weak lensing distortions of background galaxies by clusters can be used to provide accurate cluster mass estimates (e.g. Corless & King, 2009; Becker & Kravtsov, 2011; Dietrich et al., 2019; McClintock et al., 2019). However, mass measurements from weak gravitational lensing of background galaxies become extremely challenging at high redshift $z \sim 1$, where the number of background sources in typical imaging datasets drops, weakening the mass constraints. Moreover, the scatter between weak lensing inferred masses and true halo mass is large, implying that large numbers of clusters are needed for accurate mass calibration. Recently, Baxter et al. (2018) applied gravitational lensing of the Cosmic Microwave Background (CMB), using CMB maps from the South Pole Telescope (SPT) 2500 deg^2 SPT-SZ survey, demonstrating an ability to constrain the amplitude of the λ–mass relation to $\sim$20% accuracy. This offers great promise for the future, assuming systematic biases due to the thermal Sunyaev-Zel'dovich effect and cluster mis-centering can be accurately corrected. Cluster velocity dispersions, obtained through spectroscopic observations of cluster member galaxies, have proven to be good mass proxies as well, due in part to their insensitivity to complex ICM physics. But as with weak lensing masses, dispersion based masses still show large per-cluster scatter (Evrard et al., 2008; Saro et al., 2013; Sifón et al., 2013b; Ruel et al., 2014), implying that large samples must be used for mass calibration.

In this work, we aim to calibrate the λ–mass–redshift scaling relation parameters by performing a dynamical analysis based on the Jeans equation (Binney & Tremaine, 1987). In particular, we use a modification of the MAMPOSSt technique (Modeling Anisotropy and Mass Profiles of Observed Spherical Systems; Mamon et al., 2013), which fits the distribution of particles in the observed projected phase space (line of sight velocities and distribution as a function of projected radius), to use the full information in the LOS velocity distribution and projected positions of cluster galaxies. This method has been extensively used to recover dynamical masses and gain information on galaxy formation and evolution (e.g. Biviano et al., 2013, 2017; Munari et al., 2014). In particular, in Capasso et al. (2019) it was demonstrated that, using this method on a composite cluster with $\sim$600 cluster members, dynamical masses and orbital anisotropy of the galaxy population can be simultaneously constrained, delivering masses with a $\sim$15% uncertainty (decreasing to $\sim$ 8% when using a composite cluster with $\sim$ 3000 tracers). In addition, it was shown that combining cluster dynamical constraints in likelihood space produces final mass constraints that are consistent with masses from composite or stacked cluster analyses.

We perform a dynamical analysis on the ROSAT All-Sky Survey (RASS) X-ray cluster candidates, which have optical counterparts in SDSS imaging data identified using the redMaPPer algorithm (the red sequence Matched-filter Probabilistic Percolation algorithm,

Rykoff et al., 2014, see Section 4.2.1). The resulting cluster catalog is called CODEX (COnstrain Dark Energy with X-ray clusters; Finoguenov, in prep), and a subset of these clusters have since been spectroscopically studied within the SPectroscopic IDentification of eRosita Sources (SPIDERS) survey (Clerc et al., 2016). The analysis carried out here includes a sample of 428 CODEX clusters with a corresponding sample of ∼7800 red member galaxies with measured redshifts. The clusters span the redshift range $0.03 \leq z_c \leq 0.66$, with richness $20 \leq \lambda \leq 230$.

The paper is organized as follows: In Section 4.2 we summarize the dataset used for our analysis. In Section 4.3 we give an overview of the theoretical framework. The results are presented in Section 4.4, where we discuss the outcome of our mass–observable relation calibration, and we present our conclusions in Section 4.5. Throughout this paper we adopt a flat ΛCDM cosmology with a Hubble constant $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and a matter density parameter $\Omega_M = 0.3$. Cluster masses (M_{200c}) are defined within r_{200c}, the radius within which the cluster overdensity is 200 times the critical density of the Universe at the cluster redshift. We refer to r_{200c} simply as the virial radius. All quoted uncertainties are equivalent to Gaussian 1σ confidence regions unless otherwise stated.

4.2 Data

This work is based on a spectroscopic galaxy sample constructed within the SPIDERS survey (Clerc et al., 2016), which observed a subset of CODEX galaxy clusters. These clusters were selected from the ROSAT All-Sky Survey (RASS, see Voges et al., 1999) and then cross–matched with nearby optically selected systems identified using the redMaPPer algorithm applied to the Sloan Digital Sky Survey IV (SDSS-IV, see Dawson et al., 2016; Blanton et al., 2017) optical imaging data. In the following section we describe each of these elements of the dataset.

4.2.1 The redMaPPer algorithm

RedMaPPer is an optical cluster-finding algorithm based on the red sequence technique, built around the richness estimator of Rykoff et al. (2012). It has been successfully applied to photometric data from the Eighth Data Release (DR8; Aihara et al., 2011) of the Sloan Digital Sky Survey, and subsequently to the SDSS Stripe 82 coadd data (Annis et al., 2014) and to the Dark Energy Survey (DES) Science Verification Data (SV) and Year 1 (Y1) data (Saro et al., 2015; Rykoff et al., 2016; Soergel et al., 2016). It has been shown to provide excellent photometric redshift performance and optical richness estimates λ that tightly correlate with external mass proxies.

The optical catalog construction is performed in several steps. First of all, the red sequence model is calibrated on a set of clusters having spectroscopic redshifts. This model is then used to identify galaxy clusters and measure their richness. To each galaxy in the vicinity of a galaxy cluster, redMaPPer estimates the membership probability, $P_{mem} \in [0, 1]$, based on its magnitude, colors and clustercentric distance. This probability is also

used to estimate the richness of the cluster. The latter is thus defined as the sum of the membership probabilities (P_{mem}) over all galaxies $\lambda = \sum P_{mem}$.

4.2.2 The CODEX sample

The CODEX survey is designed to combine ROSAT X-ray cluster candidates with optical selected cluster candidates identified using redMaPPer (the red sequence Matched-filter Probabilistic Percolation algorithm, Rykoff et al., 2014, see Section 4.2.1). This catalog is constructed in several steps. As a first step, RASS data are used to identify all X-ray sources with detection significance S/N$>$ 4. The redMaPPer algorithm is then run on the SDSS imaging data around each RASS source position to identify candidate clusters with a red sequence, which constitutes a collection of passive galaxies lying at a common redshift. The redMaPPer algorithm provides an estimate for the photometric redshift of the cluster, an estimation of the optical richness and an optical cluster center, which is constrained to be within $3'$ of the X-ray position. In cases of multiple optical counterparts meeting these criteria, the counterpart having the highest richness is assigned to the RASS X-ray source.

Using the updated optical position of the cluster, a revised red sequence is identified, providing the final estimate of the cluster photometric redshift and richness (optical or "OPT" quantities: $z_{\lambda,\mathrm{OPT}}, \lambda_{\mathrm{OPT}}$, etc.). If the cluster is at sufficiently high redshift that the SDSS photometric data are not deep enough to allow a direct measurement of richness over a fixed fraction of the cluster galaxy luminosity function (i.e., to a limit $m_*(z) + \Delta$, where Δ is the same for all clusters), then a correction factor η is calculated and applied to the richness. As described in Section 4.4.1, this has an impact on the Poisson noise contribution to the richness and must be included in the analysis of the mass–observable scaling relation.

In the final step, X-ray properties based on the RASS count-rate and the redMaPPer redshift are calculated in optimized apertures (imposing a minimal signal-to-noise threshold of 1.6), assuming a model for the X-ray spectral emissivity, along with the aperture-corrected cluster flux f_X and [0.1-2.4] keV luminosities L_X. The final CODEX sample then results in X-ray detected clusters, for which we have an estimate of the redshift, optical richness, the optical cluster center, and X-ray luminosity. This sample has been used for follow-up observations of the SPIDERS survey, described below, which finally provided spectroscopic redshift estimates of cluster member galaxies.

4.2.3 The SPIDERS spectroscopic sample

The SPIDERS survey is an observational program aiming to obtain homogeneous and complete spectroscopic follow-up of extragalactic sources, using data from X-ray satellites that lie within the SDSS-IV imaging footprint. The driving goals of the program are the confirmation of X-ray extended sources identified as galaxy cluster candidates and the assignment of a precise redshift. In the final years of SDSS-IV, SPIDERS will follow-up X-ray extended sources extracted from the all sky X-ray eROSITA survey (extended ROentgen Survey with an Imaging Telescope Array Predehl et al., 2010; Merloni et al., 2012). Prior

to the launch of eROSITA, galaxy clusters identified in the shallower RASS and sparser XMM-Newton data will constitute the bulk of the SPIDERS program. The spectroscopy is obtained using the BOSS spectrograph mounted on the SDSS-2.5m telescope at Apache Point Observatory (Gunn et al., 2006), performing follow-up of galaxies detected in the large area of extragalactic sky imaged in *ugriz* filters by the same telescope. In the following sections we describe the target selection, the cuts made on the sample, and how the spectroscopic galaxy sample used in this work is obtained.

Target selection

The target selection and initial cuts to the sample are outlined in Clerc et al. (2016). Here, we summarize the most salient features. To optimize the number of spectroscopically confirmed clusters, the redMaPPer membership probability $P_{\rm mem}$ is used as a reference to assign priorities to potential targets, ranking galaxies within each cluster. The algorithm starts with the richest cluster in the sample, iteratively proceeding to lower richness. The pool of targets along with the priority flag is then submitted to the eBOSS tiling algorithm. The final data reduction and spectral classification relies on the eBOSS spectroscopic pipeline and processing.

An automatic procedure is used to assign the membership of red sequence galaxies with measured redshifts. For each cluster, an iterative clipping procedure is performed. As a first step, members with velocity offsets greater than 5000 km/s (relative to this first guess mean redshift) are rejected. The remaining potential members $N_{\rm z\text{-}spec}$ are used to estimate the velocity dispersion of the cluster, either using the bi-weight variance ($N_{\rm z\text{-}spec} \geq 15$; see Beers et al., 1990) or the gapper estimator (if $N_{\rm z\text{-}spec} < 15$). Finally, a 3σ clipping is applied, rejecting objects lying further away than 3 times the velocity dispersion from the mean velocity.

A final validation of all galaxy clusters and assessment of their redshifts is achieved through visual screening of the outcome of the automatic procedure. Sometimes the automated procedure fails. This occurs, for example, if fewer than 3 members are assigned to a cluster, or if the initial 5000 km/s clipping rejected all members. The latter can occur when there are several distinct structures along the line of sight. Independent inspectors analyze these complex cases, which may lead to inclusion or removal of members. This process sets the validation status and mean redshift of the cluster. Line-of-sight projection effects not disentangled by the photometric membership algorithm can also be identified and split into several components. Final cluster redshift estimates are based on the bi-weight average (Beers et al., 1990) of all red sequence galaxies selected as cluster members, if at least 3 members are assigned to the cluster. The typical cluster redshift statistical uncertainty is $\Delta_z/(1+z) \lesssim 10^{-3}$.

The updated cluster spectroscopic redshifts are then used to update the computation of X-ray cluster properties. Assuming the standard flat ΛCDM cosmological model (Hubble constant $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and matter density parameter $\Omega_{\rm M} = 0.3$, ROSAT fluxes are converted into rest-frame [0.1-2.4] keV luminosities and scaling relations allow an estimate of the cluster mass and characteristic radius r_{500} or r_{200c}. The typical measurement

uncertainty on the luminosities of CODEX clusters amounts to $\approx 35\%$, as computed from the Poissonian fluctuation in the associated ROSAT X-ray photons (see Mirkazemi et al., 2015).

Final spectroscopic cluster member sample

Given the sample produced as described above, we apply some additional cuts prior to our analysis. As mentioned above, there are cases in which a CODEX cluster has multiple groups of galaxies separated by a large velocity gap along the line of sight. To avoid merging systems, we only use clusters which are flagged as having one component along the line of sight. We restrict our analysis to the cluster virial region ($R \leq r_{200c}$). Moreover, we exclude the very central cluster region ($R \leq 50$kpc), to account for the positional uncertainties of cluster centers, and to avoid including the centrally located BCG in the dynamical analysis. At the end of this process, our spectroscopic dataset from SPIDERS consists of 705 galaxy clusters, for a total of ≈ 11400 candidate cluster members, with a median redshift $z = 0.21$ and spanning a richness range $20 \leq \lambda \leq 230$. At the time this paper is being written, the observations of the galaxy clusters included in our sample have already been completed. No further galaxy spectroscopic redshifts will be assigned to these clusters during the final stages of the SDSS-IV program.

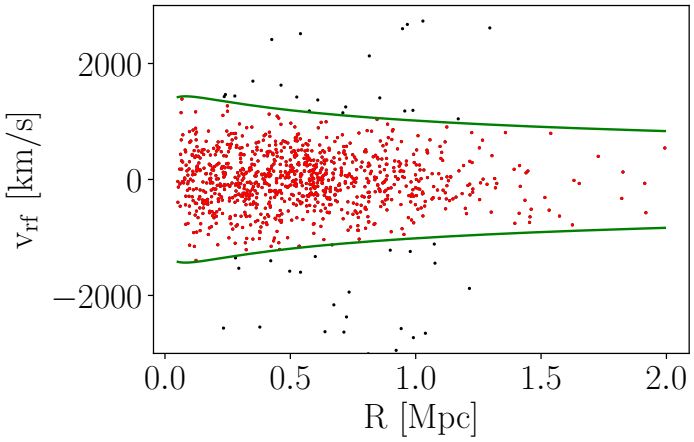

Figure 4.1: The projected phase space diagram for the composite cluster constructed using those objects having richness in the range $20 \leq \lambda \leq 23.5$. Green lines represent the radially-dependent $2.7\sigma_{\rm LOS}$ cut used to reject interlopers (indicated by black dots).

4.2.4 Interloper rejection

The observables on which the analysis is based are the galaxy projected clustercentric distance R and the rest frame line of sight (LOS) velocity v_{rf}. We extract v_{rf} from the galaxy redshift z_{gal} and equivalent velocity $v(z_{gal})$ as $v_{rf} \equiv (v(z_{gal}) - v(z_c))/(1 + z_c)$, with z_c being the cluster redshift.

Even though the SPIDERS automated procedure assesses membership for each galaxy, there could still be interloper galaxies, i.e. galaxies that are projected inside the cluster virial region, but do not actually lie inside it. To reduce this contamination, we apply the "Clean" method (Mamon et al., 2013), which uses the projected phase space location of each galaxy and its comparison to the expected maximal line of sight velocity at each projected radius estimated for the cluster. Because we do not have enough spectroscopic redshifts to do this accurately for each individual cluster, we divide our sample in bins of richness and perform the interloper rejection in each of them separately. Specifically, we divide the sample into 15 equally spaced λ bins and build a composite cluster in each bin. We apply no scaling in velocity, and stack in physical radius [Mpc] to build the composite clusters.

The "Clean" method is implemented through several steps. First, the cluster mass is estimated from the LOS velocity dispersion $\sigma_{\rm LOS}$ of each composite cluster, using a scaling relation calibrated using numerical simulations (e.g., Saro et al., 2013), and assuming an NFW mass profile with concentration sampled from the theoretical mass–concentration relation of Macciò et al. (2008b). Thereafter, assuming the Mamon & Łokas (2005, ML) velocity anisotropy profile model and given the $M(r)$ of the cluster, a Gaussian LOS velocity dispersion profile with $\sigma_{\rm LOS}(R)$ is calculated and used to iteratively reject galaxies with $|{\rm v_{rf}}| > 2.7\sigma_{\rm LOS}$ at any clustercentric distance (see Mamon et al., 2010a, 2013). In Fig. 4.1 we show the location of galaxies in projected phase space with the identification of cluster member galaxies for the composite cluster constructed using those objects having richness in the range $20 \leq \lambda \leq 23.5$.

The distribution of the final sample of galaxies in projected phase space is presented in Fig. 4.2. In this plot we show the galaxies identified as cluster members (red dots), the rejected interlopers (black dots), and the radial and velocity distributions of the member galaxies with measured redshifts (green histograms).

We also note that, even after carrying out interloper rejection, there is still a degree of contamination by interlopers. In fact, galaxies that lie outside the virial radius will tend to have smaller peculiar velocities than those galaxies lying within the virial region. Indeed, close to the cluster turn-around radius the galaxies will have negligible peculiar velocity and cannot be removed from the sample through an interloper rejection algorithm of the type we adopt here. In fact there is no obvious method for separating these galaxies from the sample within the cluster virial region that we wish to model. An analysis of cosmological N-body simulations carried out by Saro et al. (2013) shows that, when passive galaxies are selected, this contamination is characteristically $\sim$20% for massive clusters ($M_{\rm 200c} \geq 10^{14} M_\odot$). For less massive clusters the contamination is expected to be higher. Another work carried out by Mamon et al. (2010a), based on hydrodynamical cosmological simulations, showed that the distribution of interlopers in projected phase space is nearly universal, presenting only small trends with cluster mass. They state that, even after applying the iterative $2.7\sigma_{\rm LOS}$ velocity cut, the fraction of interlopers is still 23 ± 1% of all DM particles with projected radii within the virial radius, and over 60% between 0.8 and 1 virial radius. Further exploration of the effects of this contamination on the dynamical analysis is required, and we are pursuing that in a separate study (Capasso

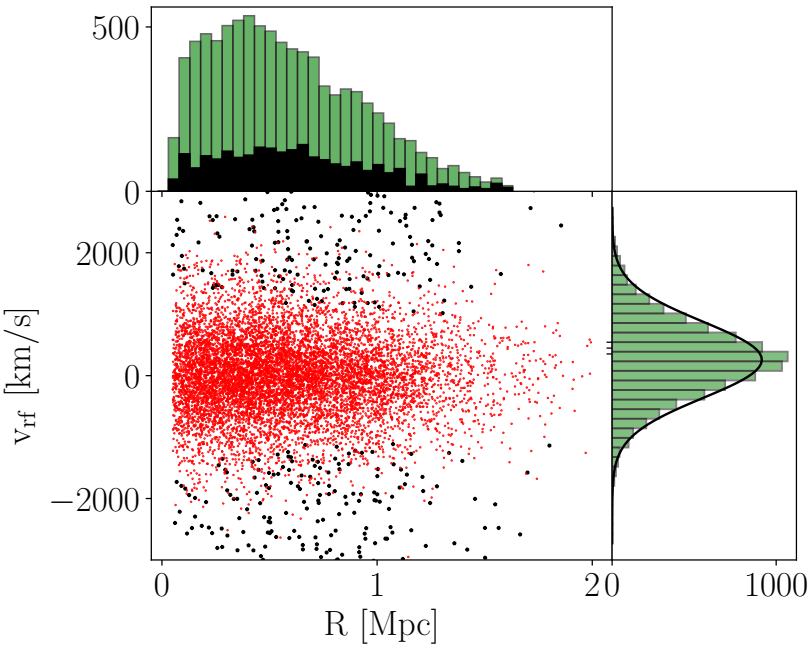

Figure 4.2: Projected phase space distribution for the final sample of 428 clusters. Red dots indicate the 7807 cluster members, while black dots mark the ∼2000 rejected interloper galaxies. In the upper panel we show in green the radial distribution of the member galaxies with measured redshifts, and in black the radial distribution of the interlopers. The panel on the right shows the distribution of rest-frame velocities, with an overplotted Gaussian of the same dispersion for comparison.

et al., in prep.).

After the application of the interloper rejection, we are left with a total of 703 clusters and 9121 red galaxies. For the analysis presented here we apply another cut on the cluster sample, keeping all CODEX systems that currently have at least 10 spectroscopic members, $N_{\rm mem} \geq 10$. After this cut, our sample consists of 428 clusters and 7807 red galaxies, with a median redshift, richness, and luminosity of z = 0.18, λ=41, and $L_X = 9.2 \times 10^{43} {\rm erg\,s^{-1}}$, respectively. Fig. 4.3 shows the distributions of cluster redshift and richness of the final sample.

4.2.5 Galaxy number density profile

The Jeans analysis requires knowledge of the 3D number density profile $\nu(r)$ of the tracer populations whose dynamical properties are being used to study the mass and orbital properties of the system. In our case, these are the red sequence galaxies selected by the redMaPPer algorithm for observations within SPIDERS. As only the logarithmic derivative of $\nu(r)$ enters the Jeans equation (see equation 4.1), the absolute normalization of the galaxy number density profile has no impact on our analysis. However, a radially dependent incompleteness in the velocity sample would impact our analysis. In general, the spectro-

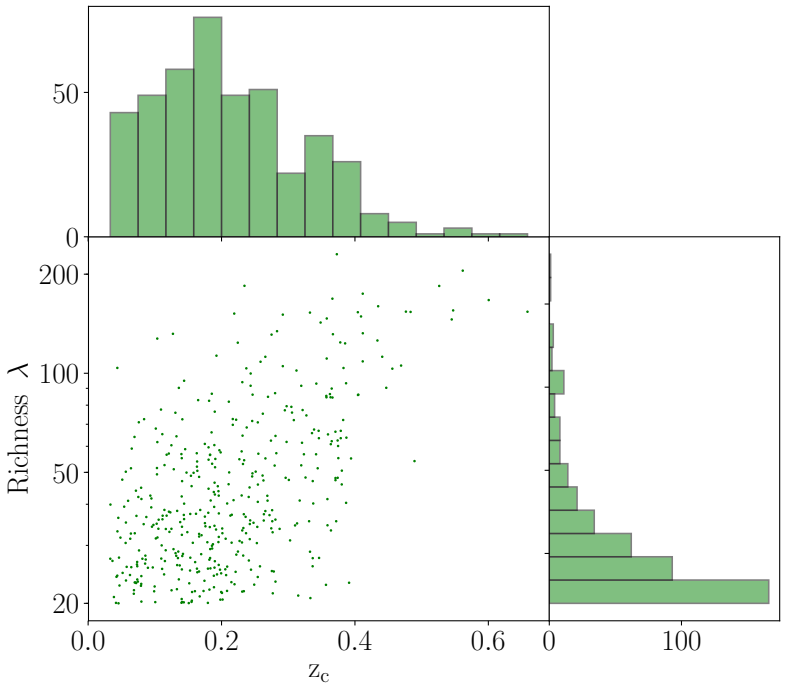

Figure 4.3: Distribution of richness λ and cluster redshift z_c of the final cluster sample.

scopic followup within SPIDERS will lead to a radially dependent incompleteness. This means we cannot simply adopt the spectroscopic sample to measure the number density profile of the tracer population. We therefore rely on a study of the galaxy populations in 74 Sunyaev-Zel'dovich effect (SZE) selected clusters from the SPT-SZ survey, which have been imaged as part of the Dark Energy Survey Science Verification phase (Hennig et al., 2017). That study found no mass or redshift trends in the radial distribution of red sequence galaxies for $z > 0.25$ and $M_{200c} > 4 \times 10^{14} M_{\odot}$, finding the number density profile of the red sequence population to be well fit by a Navarro, Frenk and White (NFW) model (Navarro et al., 1996) out to radii of $4r_{200c}$, with a concentration for cluster galaxies of $c_{gal} = 5.37^{+0.27}_{-0.24}$. Therefore, we adopt the number density profile described by an NFW profile with the above-mentioned value of c_{gal} and a scale radius $r_{s,\,gal} = R_{200c}/c_{gal}$, making the assumption that the dynamical properties of our spectroscopic sample are consistent with the dynamical properties of the red sequence galaxy population used to measure the radial profiles.

We note that the Hennig et al. (2017) study indicates significant cluster to cluster scatter in the NFW concentration. We do not expect this scatter to be a source of significant bias in our analysis, because in an earlier analysis Capasso et al. (2019) showed that the mean masses extracted from composite clusters and from the fitting of an ensemble of individual clusters are in good agreement. We will nevertheless further examine the impact of mismatch between the model and actual radial distribution of the tracer population in an upcoming study where we seek to improve the understanding of biases and scatter in dynamical mass estimators using mock observations of structure formation simulations

(Capasso et al., in prep.).

4.3 Theoretical Framework

The method we adopt for the dynamical analysis of our clusters is based on the spherically-symmetric Jeans analysis (Binney & Tremaine, 1987). Using the Jeans equation, it is possible to define the mass distribution $M(r)$ of a cluster as

$$\frac{GM(<r)}{r} = -\sigma_r^2 \left(\frac{d\ln\nu}{d\ln r} + \frac{d\ln\sigma_r^2}{d\ln r} + 2\beta \right), \tag{4.1}$$

where $\nu(r)$ is the number density profile of the tracer galaxy population, $\sigma_r(r)$ is the radially dependent component of the velocity dispersion along the spherical coordinate r, $M(<r)$ is the enclosed mass within radius r, G is Newton's constant, $\beta(r) \equiv 1-(\sigma_\theta^2/\sigma_r^2)$ is the radially dependent velocity dispersion anisotropy, and σ_θ is the tangential component of the velocity dispersion. The observables we employ to constrain these quantities are projected quantities, including the surface density profile of the galaxy distribution, the rest frame LOS velocities and the radial separation of each galaxy from the cluster center.

Given the limited knowledge of the line of sight velocity distribution within realistic cluster dynamical datasets, it is not possible to uniquely derive the mass distribution of a galaxy cluster (Merritt, 1987). To address this problem, we use the Modeling Anisotropy and Mass Profiles of Observed Spherical Systems algorithm (hereafter MAMPOSSt; for full details please refer to Mamon et al., 2013). This code performs a maximum likelihood analysis of the projected phase space distribution of the observed sample using the theoretical distribution predicted for a given model using the Jeans equation. The observations are used to constrain the model parameters adopted to describe the cluster mass distribution and galaxy orbital anisotropy. The MAMPOSSt method thus requires adopting parametrized models for the number density, mass, and velocity anisotropy profiles $\nu(r)$, $M(r)$, $\beta(r)$. As addressed in Section 4.2.5, because our spectroscopic dataset might suffer from radially dependent incompleteness, we adopt the measured number density profile derived from the study of red sequence galaxies in SZE selected clusters (Hennig et al., 2017). We discuss our choice of the mass and velocity anisotropy profiles in the next section.

4.3.1 Mass and anisotropy profiles

Taking guidance from both numerical studies of structure formation and observational results, we adopt the mass model introduced by Navarro et al. (1996, NFW)

$$\rho(r) = \rho_0 \left(\frac{r}{r_\mathrm{s}}\right)^{-1} \left(1+\frac{r}{r_\mathrm{s}}\right)^{-2}, \tag{4.2}$$

where ρ_0 is the central density, and r_s is the scale radius where the logarithmic derivative of the density profile reaches -2. Integrating this density profile up to $r_{200\mathrm{c}}$, we obtain the

mass enclosed inside the virial radius

$$M_{200c} = 4\pi\rho_0 r_s^3 \left[\ln\left(\frac{r_s + r_{200c}}{r_s}\right) - \frac{r_{200c}}{r_s + r_{200c}}\right]. \tag{4.3}$$

Cosmological simulations produce dark matter halos with mass profiles well described by this profile. Even though some results have preferred different models (Merritt et al., 2006; Navarro et al., 2010; Dutton & Macciò, 2014; van der Burg et al., 2015b; Sereno & Ettori, 2017), this result is in good agreement with a variety of observational analyses using both dynamics and weak lensing (Carlberg et al., 1997b; van der Marel et al., 2000b; Biviano & Girardi, 2003; Katgert et al., 2004b; Umetsu et al., 2011).

For the velocity anisotropy profile, we consider five models that have been used in previous MAMPOSSt analyses and that are described also in Capasso et al. (2019). These are (1) the constant anisotropy model (C), (2) the Tiret anisotropy profile (Tiret et al., 2007, T), (3) the Mamon & Łokas (2005) profile (ML), (4) the Osipkov-Merritt anisotropy profile (Osipkov, 1979; Merritt, 1985, OM), and (5) a model with anisotropy of opposite sign at the center and at large radii (O).

Therefore, to predict the projected phase space distribution of the observed dynamical dataset for each cluster, we run MAMPOSSt with 3 free parameters: the virial radius r_{200c}, the scale radius r_s of the mass distribution, and a velocity anisotropy parameter θ_β. This parameter represents the usual $\beta = 1 - (\sigma_\theta^2/\sigma_r^2)$ for the first three models (C, T, O), while for the ML and OM models it defines a characteristic radius $\theta_\beta = r_\beta$.

4.3.2 Bayesian model averaging

As the literature does not provide us with strong predictions for the radial form of the velocity anisotropy profile $\beta(r)$, we employ all the five models described above when estimating the cluster masses. We combine the results from the different models by merging their constraints exploiting the Bayesian model averaging technique. A weight is assigned to each model, which is proportional to how well the model fits the data. This weight is represented by the so-called Bayes factor (see Hoeting et al., 1999, and references therein).

Considering the 5 anisotropy models M_1, ..., M_5, we define the Bayes factor B_j of each model j by normalizing the marginalized likelihood of the model $\mathcal{L}(D\,|M_j)$, also known as evidence, by the likelihood of the most probable model. Specifically,

$$B_j = \frac{\mathcal{L}(D\,|M_j)}{\mathcal{L}(D\,|M_{\max})}, \tag{4.4}$$

where $M_{\max}$ indicates the model with the highest marginalized likelihood, $\mathcal{L}(D\,|M_j) = \int \mathcal{L}(D\,|\theta_j, M_j) P(\theta_j\,|M_j)\, d\theta_j$, $\mathcal{L}(D\,|\theta_j, M_j)$ is the likelihood of the data D given the model parameters θ_j, and $P(\theta_j\,|M_j)$ is the prior.

The average posterior distribution on the parameter common to all anisotropy models is then simply given by the weighted average of the posterior distributions of each model, with the Bayes factor as weight. To perform this Bayesian model averaging, we employ the multimodal nested sampling algorithm MultiNest (Feroz & Hobson, 2008; Feroz et al., 2009, 2013), which provides us with the evidence for each model.

A_λ	B_λ	γ_λ	r_β	$\sigma^{\rm int}_{\ln\lambda}$
$\mathcal{U}(20,50)$	$\mathcal{U}(0.5,2)$	$\mathcal{U}(-3,2)$	$\mathcal{U}(0.01,10)$	$\mathcal{N}(0.15,0.09^2)$

Table 4.1: Priors assumed for our analysis. $\mathcal{U}(i,j)$ refers to a uniform flat prior in the interval (i,j), while $\mathcal{N}(\mu,\sigma^2)$ indicates a Gaussian distribution with mean μ and variance σ^2.

Dynamical analyses using SPIDERS data	A_λ	B_λ	γ_λ
Baseline analysis: $\lambda \geq 20$, $N_{\rm mem} \geq 10$	$38.6^{+3.1}_{-4.1} \pm 3.9$	$0.99^{+0.06}_{-0.07} \pm 0.04$	$-1.13^{+0.32}_{-0.34} \pm 0.49$
Analysis with correlated scatter correction	$39.8^{+3.0}_{-3.8} \pm 4.0$	$0.98^{+0.07}_{-0.07} \pm 0.04$	$-1.08^{+0.31}_{-0.34} \pm 0.49$
Previously published results	$\lambda(3\times10^{14}M_\odot, 0.18)$	$M^{B_\lambda}_{200c}$	$(1+z)^{\gamma_\lambda}$
WL masses using DES Y1 (McClintock et al., 2019)	43.8 ± 1.3	0.73 ± 0.03	-0.10 ± 0.10
CMB WL masses (Baxter et al., 2018)	49.8 ± 10.8	0.81 ± 0.21	–
WL masses using SDSS (Simet et al., 2017)	63.1 ± 2.2	0.74 ± 0.06	–
Cluster clustering using SDSS (Baxter et al., 2016)	37.5 ± 4.4	0.84 ± 0.12	0.70 ± 0.90
Pairwise velocity dispersion with SDSS (Farahi et al., 2016)	47.7 ± 1.0	0.75 ± 0.04	–
SPT masses with RM from DES SV (Saro et al., 2015)	36.1 ± 9.1	1.16 ± 0.20	0.60 ± 0.63

Table 4.2: RedMaPPer Richness-mass-redshift scaling relation parameters from this analysis and the literature. The results from our analysis include corrections for the Eddington and Malmquist biases. Parameters are defined in equation (4.5). For results from this analysis the uncertainties are statistical, and a systematic mass uncertainty of 10% is applied to the amplitude A_λ. In the comparison to previous results, the amplitude A_λ column contains the λ at $M_{200c} = 3\times10^{14}M_\odot$ and $z = 0.18$. Conversions have been made to M_{200c} and from $E(z)$ to $(1+z)$ where needed. Note also that each of these studies was performed on a different range of mass and redshift.

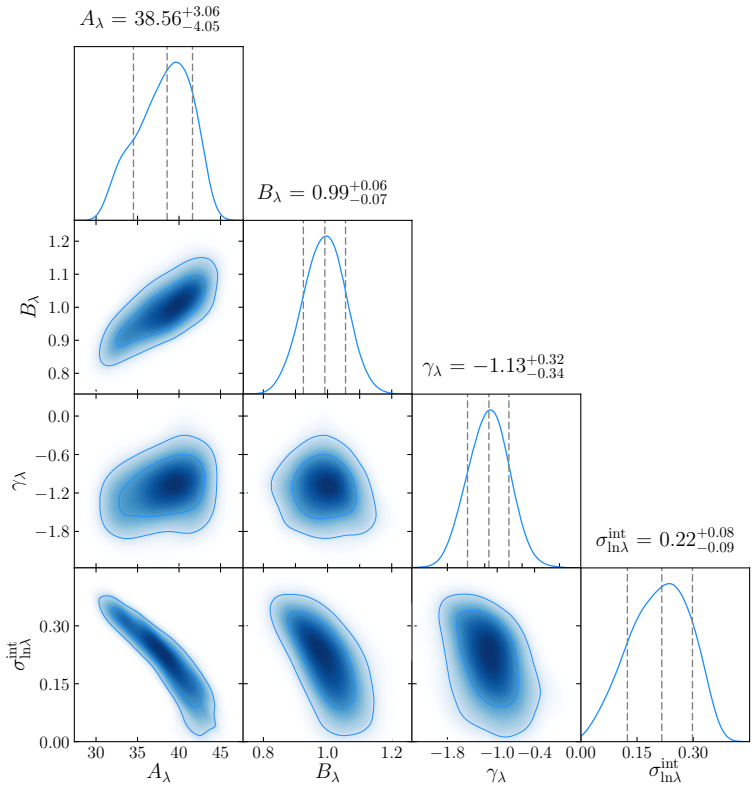

Figure 4.4: Parameters of the λ-M_{200c}-z relation. Contours show the 1σ, 2σ, and 3σ confidence regions.

4.4 Results

In this section we present the results of our dynamical analysis. In the first subsection we describe how we calibrate the λ-mass relation and present our results. In the following subsection we explore the impact of correlated scatter in the X-ray luminosity and richness for the CODEX sample. Afterwards, we compare our findings to previous works, and we test how strongly the number of member galaxies per cluster affects our results.

4.4.1 λ-M_{200c}-z relation

We adopt a power-law relation between cluster richness λ, mass and redshift of the form

$$\lambda = \mathrm{A}_\lambda \left(\frac{M_{200c}}{M_{\rm piv}}\right)^{\mathrm{B}_\lambda} \left(\frac{1+z}{1+z_{\rm piv}}\right)^{\gamma_\lambda}, \tag{4.5}$$

where A_λ, B_λ, and γ_λ are the amplitude, the mass slope and the redshift evolution slope. Similar forms have been used to study the galaxy halo occupation number and richness previously (Lin et al., 2004, 2006; Saro et al., 2015; Hennig et al., 2017; Saro et al., 2017). We adopt the redshift scaling $(1+z)^\gamma$ instead of $E(z)^\gamma$ because, as discussed in a recent study of X-ray scaling relations (Bulbul et al., 2018), we wish to avoid ascribing cosmological sensitivity to redshift trends unless there is a physically justifiable reason to do

so. Sensitivity of an observable to the evolving critical density of the Universe would justify an $E(z)$ scaling. An example would be an observable like the X-ray luminosity or Sunyaev-Zel'dovich effect signature that depends on the intracluster medium density, but in the case of the galaxy richness or halo occupation number, the density plays no role and no such sensitivity is expected. We set the pivot redshift to be $z_{\rm piv} = 0.18$, which is the median redshift of our cluster sample. We have adjusted the value of the mass pivot $M_{\rm piv} = 3 \times 10^{14} M_{\odot}$ iteratively to minimize the false degeneracy between A_λ and B_λ.

We marginalize over the intrinsic scatter in λ at fixed mass, which is set to be log-normal with a prior on the scatter from Saro et al. (2017), $\sigma^{\rm int}_{\ln\lambda} = 0.15 \pm 0.09$ (precise priors listed in Table 4.1). We assume the full scatter in λ at fixed mass is log-normal with variance given by:

$$\sigma^2_{\ln\lambda} = \frac{\eta}{\lambda} + {\sigma^{\rm int}_{\ln\lambda}}^2, \tag{4.6}$$

where η is the scale factor described in Section 4.2.2 that is a correction factor that accounts for the limited depth of the SDSS photometry in accounting for the richness calculated over a fixed portion of the cluster galaxy luminosity function.

For each cluster i in our sample, we calculate an initial mass $M_{\rm 200c,obs}$ using the scaling relation described in equation (4.5) and the current values of the parameter vector $\boldsymbol{p}$, which contains the 4 scaling relation parameters $A_\lambda, B_\lambda, \gamma_\lambda$, and $\sigma^{\rm int}_{\ln\lambda}$ together with the anisotropy model parameter r_β. In each iteration we use the current value of the scatter $\sigma_{\ln\lambda}$ to estimate a correction for the Eddington bias caused by the interplay of the λ scatter and the mass function using the method described in Mortonson et al. (2011). Assuming a log-normal mass observable relation with variance $\sigma^2_{\ln M} = (1/B_\lambda \cdot \sigma_{\ln\lambda})^2$ that is small compared with the scale over which the local slope Γ of the mass function changes, the posterior mass distribution is a log-normal of the same variance $\sigma^2_{\ln M}$ with a shifted mean $\ln < M_{\rm 200c,true} >= \ln < M_{\rm 200c,obs} > + \Gamma\sigma^2_{\ln M}$.

With this mass, we then use MAMPOSSt to construct the probability distribution in projected phase space for each cluster, combining the likelihoods calculated for each member galaxy in that cluster

$$\mathcal{L}_i = \prod_{j \in gal} \mathcal{L}(R^j, v^j_{rf}, \lambda_i, z_i \mid \boldsymbol{p}), \tag{4.7}$$

where R^j and v^j_{rf} are the clustercentric radii and rest-frame velocities of the member galaxy j in the cluster i. The maximum likelihood solutions are obtained using the NEWUOA software (Powell, 2006). Flat priors are assumed for the scaling relation parameters A_λ, B_λ, and γ_λ, and for the anisotropy parameter r_β (see Table 4.1).

We combine the likelihoods for all these clusters, to then obtain the likelihood for the total sample for each set of scaling relation parameters $\boldsymbol{p}$, i.e. $\mathcal{L} = \prod_{i \in clus} \mathcal{L}_i$. This procedure must be done separately for each anisotropy profile model (see Section 4.3.1). Finally, we use Bayesian model averaging to combine the posterior parameter distributions obtained from the different anisotropy models, effectively marginalizing over the uncertainties in the orbital anisotropy.

Because we impose a cut on our observable, $\lambda \geq 20$, a correction for the Malmquist bias is also needed (Sandage, 2000). We estimate the effect of this correction by creating a large mock catalogue (~ 4400 clusters and $\sim 165,500$ member galaxies) by computing the number of expected clusters as a function of halo mass and redshift using the halo mass function (Tinker et al., 2008). We then draw a Poisson realization of the number of expected clusters, obtaining a mass selected cluster sample with $M_{200c} \geq 7 \times 10^{13}$ and $0.05 \leq z \leq 0.66$. Using the scaling relation parameters recovered from our analysis before correcting for this bias, we calculate λ for each cluster of mass M_{200c}. Scatter is added to this relation such that the assigned λ values are sampled from a Gaussian distribution having scatter given by equation (4.6). The mock sample we produce has richnesses $\lambda > 6.5$. For each cluster in our mock sample, we create a sample of member galaxies. We run MAMPOSSt on a grid of velocities and clustercentric distances, fixing the galaxy number density profile to that described in Section 4.2.5, and generating a random number of galaxies per cluster drawn from the distribution of member galaxies in our observed sample. Finally, we use the MAMPOSSt likelihood to recover the probability density of observing an object at a certain location in phase space (see equation 11, Mamon et al., 2013).

We fit this mock dataset and recover best fit parameter values that are consistent with the input values. Then we impose a $\lambda > 20$ richness cut on the sample and refit, noting that the best fit mass and redshift trends are affected. Using this approach, we estimate corrections for the Malmquist bias that correspond to δB_λ=+0.05 and $\delta\gamma_\lambda$=-0.06. These corrections are included in all the results we present.

Table 4.2 summarizes the posterior of our model parameters from our so-called "baseline analysis", i.e. before accounting for the impact of correlated scatter (see Section 4.4.2), while Fig. 4.4 shows the corresponding joint parameter constraints. Our results imply that galaxy clusters with redshift $z = 0.18$ and mass $M_{200c} = 3 \times 10^{14} M_\odot$ have a mean richness of $A_\lambda = 38.56^{+3.06}_{-4.05}$. The mass scaling is consistent with linear, $B_\lambda = 0.99^{+0.06}_{-0.07}$. The redshift dependence in the CODEX sample is $\gamma_\lambda = -1.13^{+0.32}_{-0.34}$, indicating that the red sequence richness λ at fixed mass falls as one moves to higher redshift.

4.4.2 Additional Systematic Effects

The results presented in the last section include corrections for the Eddington bias and the Malmquist bias, but the uncertainties on the parameters reflect only statistical errors. In this section we consider systematic effects and the impact they have on the best fit parameters and the parameter uncertainties.

We estimate that there is an additional 10% systematic uncertainty associated with the dynamical mass measurements themselves. This estimate comes from an analysis of the MAMPOSSt code run on numerical simulations in the analysis of Mamon et al. (2013). In their work, the authors show that, using particles lying within a sphere of r_{100} around the halo center, the estimate of the cluster virial radius r_{200c} is biased at $\leq 3.3\%$ (see Table 2, Mamon et al., 2013). Therefore, we adopt a Gaussian systematic uncertainty on the virial mass M_{200c} of $\sigma = 10\%$. The Mamon et al. (2013) analysis does not explore

Number of cluster member galaxies	A_λ	B_λ	γ_λ
$N_{\rm mem} \geq 1$	$39.2^{+2.9}_{-3.5}$	$0.91^{+0.05}_{-0.06}$	$-0.15^{+0.23}_{-0.24}$
$N_{\rm mem} \geq 3$	$39.3^{+3.1}_{-3.6}$	$0.92^{+0.05}_{-0.06}$	$-0.26^{+0.23}_{-0.24}$
$N_{\rm mem} \geq 5$	$39.2^{+3.0}_{-3.7}$	$0.95^{+0.06}_{-0.06}$	$-0.65^{+0.26}_{-0.27}$
$N_{\rm mem} \geq 10$	$38.6^{+3.1}_{-4.1}$	$0.99^{+0.06}_{-0.07}$	$-1.13^{+0.32}_{-0.34}$
$N_{\rm mem} \geq 20$	$41.6^{+2.5}_{-3.2}$	$0.98^{+0.09}_{-0.08}$	$-1.00^{+0.49}_{-0.56}$

Table 4.3: Impact of the number of spectroscopic members on the RedMaPPer Richness-mass-redshift scaling relation parameters. Parameters are defined in equation (4.5). The uncertainties on the results are statistical, corresponding to 68 per cent confidence intervals, and a systematic mass uncertainty of 10% is applied to the amplitude A_λ.

mass or redshift trends in these biases, and therefore we apply the entire uncertainty to the normalization parameter A_λ. In a future analysis, we plan to explore the mass and redshift dependence of the systematic uncertainties in dynamical mass estimates from a Jeans analysis (Capasso et al., in prep.).

In the subsections below we first consider the impact of selecting different subsamples using the number of member galaxies with spectroscopic redshifts $N_{\rm mem}$, and then we explore the impact of possible correlated optical and X-ray scatter.

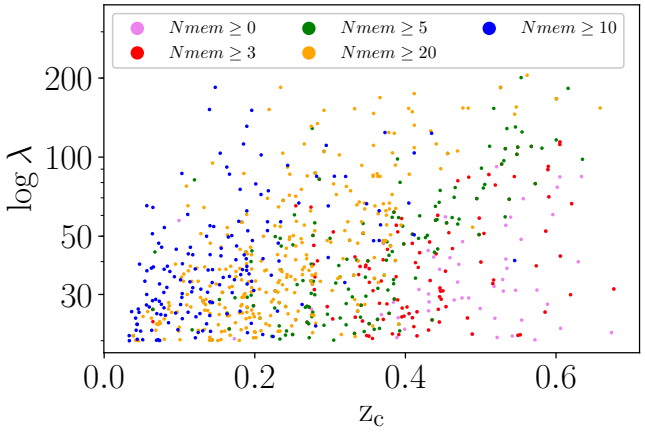

Figure 4.5: Richness and redshift distribution of clusters having a different number of spectroscopic members.

Impact of number of cluster member galaxies

As described in Section 4.2.3, we apply a cut to our sample prior to the dynamical analysis, keeping only those systems having at least 10 spectroscopic members: $N_{\rm mem} \geq 10$. This decision is driven by our concern that good constraints on the cluster masses and scaling relation parameters could not be obtained from clusters having very small numbers of

spectroscopic members. However, this selection is somewhat arbitrary, and so we explore here the impact of varying this cut.

Table 4.3 shows the results obtained imposing different cuts on the number of spectroscopic members, where $N_{\rm mem}$ varies from 1 to 20. Note that the BCG has been excluded, so the clusters with a single galaxy actually have two measured redshifts. Interestingly, the normalization A_λ and the mass trend parameter B_λ are not significantly affected when analyzing clusters having a different number of spectroscopic members.

On the other hand, the value of the redshift trend parameter γ_λ varies considerably, even reaching values consistent with zero evolution when including clusters having $N_{\rm mem} \geq 1$ and $N_{\rm mem} \geq 3$. The value of γ_λ becomes stable when including only clusters with at least 10 spectroscopic members, justifying our decision of including only those clusters into our main analysis. However the strong dependence of γ_λ on the member cut is an indication of remaining systematic uncertainties on this parameter.

The reason of the different behavior of γ_λ with respect to that of A_λ and B_λ is clarified to some degree in Fig. 4.5, where we show the distribution in richness and redshift of galaxy clusters having a different number of spectroscopic members. The distribution of clusters having $N_{\rm mem} < 10$ extends to higher redshifts, allowing for improved constraints on the redshift trend and also introducing a qualitatively different population of clusters into the analysis.

As the spectroscopic sample at these higher redshifts is increased, we will begin to see whether the trend in γ_λ with the $N_{\rm mem}$ cut is revealing a systematic in dynamical masses in the limit of very low spectroscopic sampling of each halo or whether the weaker trends shown with the less dramatic cuts that then include more high redshift systems is really a reflection of the true redshift trend in the λ-mass relation. But at this point we use the trend in γ_λ that is apparent in Table 4.3 to estimate a systematic uncertainty on that parameter. Specifically, we adopt half the full range of variation in the value as the systematic uncertainty on the parameter $\sigma_{{\rm sys},\gamma_\lambda} = \frac{\Delta|\gamma_\lambda|}{2} = 0.49$. Similarly for the mass trend parameter we estimate $\sigma_{{\rm sys},B_\lambda} = \frac{\Delta|B_\lambda|}{2} = 0.035$. For the amplitude parameter A_λ the shift is small compared to the 10% systematic uncertainty described at the beginning of this section. These systematic uncertainties are listed in Table 4.2.

Impact of correlated λ and $L_{\rm X}$ scatter

Before comparing our results to those from the literature, we examine the impact of correlated scatter in the richness and X-ray luminosity on the parameters of the richness mass relation. To do this we employ the selection function of the CODEX survey calculated as described below by the CODEX team.

As described above in Section 4.2.2, the CODEX cluster catalog is based on the identification of faint X-ray sources with the help of redMaPPER follow-up on the SDSS photometry to identify optical counterparts. The final catalog is therefore subject to both X-ray and optical selection in a manner that has been modeled based upon several observational results. First, the LoCuSS survey (Local Cluster Substructure Survey Okabe et al., 2010; Haines et al., 2018) indicates a negative value of the covariance at fixed mass of the

scatter in the X-ray luminosity L_X and the optical richness λ. For the selection function modeling adopted here, the covariance coefficient is fixed to be $\rho_{L_X-\lambda} = -0.2$ (Farahi et al. submitted). The net effect of this correlated scatter is that the CODEX survey is more sensitive in detecting clusters of given mass if they have lower richness, because that lower richness is correlated to a higher X-ray luminosity. The modeling of the survey selection function takes into account the covariance of the scatter in L_X-mass relation with the shape of the cluster, which affects the sensitivity to a particular cluster. In modeling the selection function, the scaling relations are fixed to those of the XXL survey (e.g. Pacaud et al., 2016), which is well suited for our study here, because it includes both cluster and group mass scales.

Using the selection function described above, the CODEX team then estimated the effective solid angle of the CODEX survey as a function of the scatter in λ as a function of redshift and mass. The idea here is that because scatter to lower λ is weakly correlated to an increase in the cluster X-ray luminosity, one is effectively probing a larger solid angle for those clusters with lower than typical λ at each redshift and mass. It is with this data product that we begin our analysis.

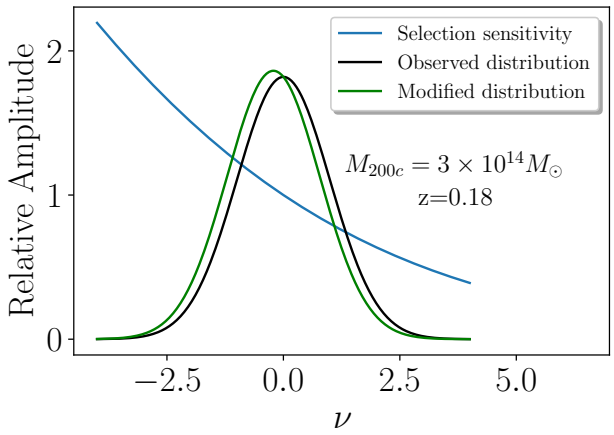

Figure 4.6: Effect of the selection function on the λ distribution. In blue we show the the relative sensitivity of the CODEX X-ray selected sample as a function of deviation from the mean observed λ, normalized to its value at $\nu = 0$. The black curve shows the distribution of observed λ, as a function of deviation ν from the mean value, while the green distribution shows how the inclusion of the selection sensitivity causes a shift and distortion of the observed λ distribution.

To estimate the impact of this correlated scatter on our results, we calculate its effects *a posteriori*, using the results of our baseline analysis as listed in Table 4.2. The variation in sensitivity as a function of λ at fixed mass and redshift produces a modification in the shape of the richness distribution at each mass and redshift. In Fig. 4.6 we show an example of how this affects the cluster distribution in λ at $M_{200c} = 3 \times 10^{14} M_\odot$ and $z = 0.18$. The blue line represents the relative sensitivity $s(\nu)$ of the CODEX X-ray selected sample as a function of the deviation ν from the mean, expected λ (expressed in equation 4.5). This

deviation is defined as a function of

$$\nu = \frac{\Delta \ln \lambda}{\sigma^{\rm int}_{\ln \lambda}}, \tag{4.8}$$

and the sensitivity function has been normalized to its value at $\nu = 0$. The black curve shows the log-normal parent distribution of λ at this mass and redshift (equation 4.6), as a function of the deviate ν. In this space this distribution is simply a Gaussian of unit width. The green distribution shows the product of the parent λ distribution with the selection sensitivity. Given the ν dependence of the sensitivity, the new λ distribution is well approximated as being a new log-normal distribution with mean shifted away from zero. The shift in the parent λ distribution can be written

$$\langle \nu | M_{200c}, z \rangle = \int d\nu \, P(\nu) \, s \left(\nu | M_{200c}, z \right). \tag{4.9}$$

where $P(\nu)$ is the parent λ distribution (log-normal) and $s\left(\nu|M_{200c}, z\right)$ is the sensitivity as a function of ν given the cluster mass and redshift. For the given example, the mean shift is $\langle \nu | M_{200c}, z \rangle = -0.20$. This shift changes little with mass, but it does evolve with redshift. This fractional logarithmic shift then implies a shift in λ for any given mass and redshift

$$\lambda_{\rm cor} = \lambda \left(M_{200c}, z \right) e^{-\langle \nu | M_{200c}, z \rangle \sigma_{\ln \lambda}} \tag{4.10}$$

To estimate the impact on the scaling relation parameters, we calculate $\lambda_{\rm cor}$ over the full range of M_{200c}, z where we have clusters. Using these results, we fit a scaling relation of the same form as equation (4.5) to the corrected data. Table 4.2 contains the best fit parameters and one sigma uncertainties of the λ– mass relation with the correlated scatter correction. The impact of the correlated scatter in λ and $L_{\rm X}$ is smaller than the 1σ statistical parameter uncertainty for all three parameters. Thus, for a sample the size of the current SPIDERS analysis, this effect can be safely ignored.

4.4.3 Comparison to previous results

In this section we compare our calibration of the richness-mass relation to previous results from the literature. We show the mass and redshift trends of λ in Figs. 4.7 and 4.8, respectively, where for the redshift trend we correct the data points to the mass $M_{200c} = 3 \times 10^{14} M_\odot$ and for the mass trend we move the data points to the redshift $z = 0.18$. These are the mass and redshift pivots of our sample, and are therefore the places where our constraints are tightest. The best fit model for the $\lambda - M_{200c}$ relation is shown in red, with shaded 1σ and 2σ confidence regions. For the results from Saro et al. (2015, in cyan) and McClintock et al. (2019, in blue), we show only the 1σ confidence region. We limit the redshift range to that analyzed in each work. Fig. 4.7 makes clear that the mass slope of our relation lies in between that of Saro et al. (2015) and McClintock et al. (2019). In Fig. 4.8 our results suggest stronger negative redshift evolution than in either previous results.

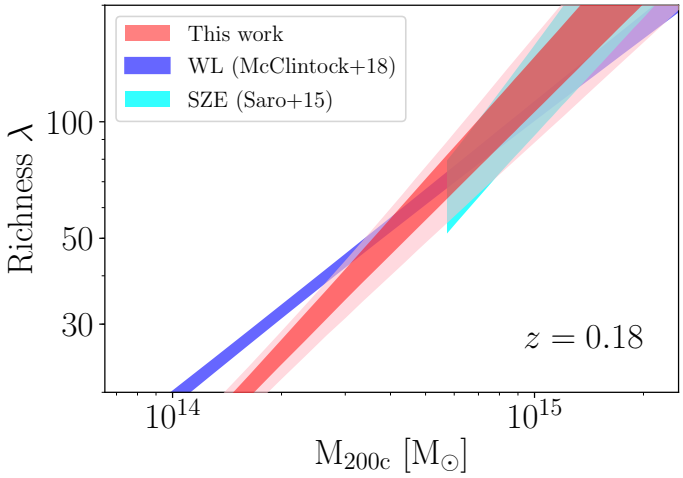

Figure 4.7: Best fit model for our richness-mass relation (in red), evaluated at the redshift $z = 0.18$, compared to other measurements. For our analysis we also show the 2σ confidence area (pink region around the red relation). Confidence regions include statistical errors only.

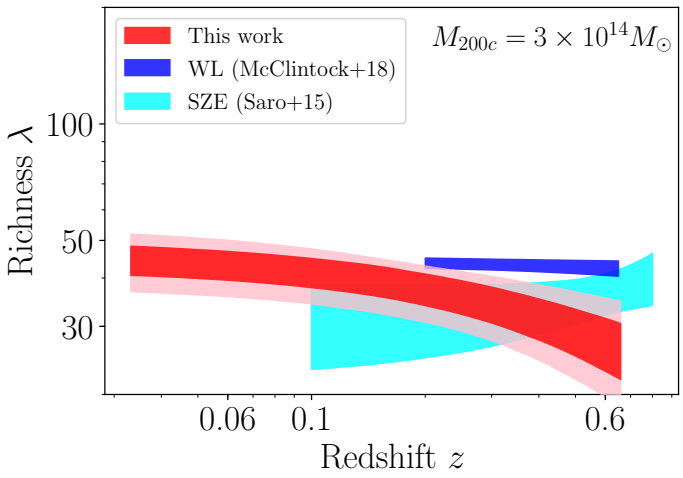

Figure 4.8: Best fit model for our richness-redshift relation (in red), evaluated at our pivot mass $M_{\rm piv} = 3 \times 10^{14} M_\odot$, compared to previous works. For our analysis we also show the 2σ confidence region. Confidence regions include statistical errors only.

Table 4.2 lists the parameter estimates and uncertainties for all the comparison results. To make these comparisons, we scale all the measurements from previous analyses to the redshift $z_{\rm piv} = 0.18$ (Fig. 4.7), and mass $M_{\rm piv} = 3 \times 10^{14} M_\odot$ (Fig. 4.8), using the best fit redshift and mass trends published for each sample. Doing this, we predict the $\lambda(3{\times}10^{14} M_\odot, 0.18)$ for each previous work. All mass conversions needed for the comparison plot are carried out using Colossus, an open-source python package for calculations related to cosmology (Diemer, 2017). The mass and redshift trend parameters presented in Table 4.2 were also converted to those defined in equation (4.5) using the appropriate mass definition $M_{\rm 200c}$ and redshift trend function $(1+z)^{\gamma_\lambda}$ adopted for our analysis here.

In some cases this involved inversions of the mass-observable relations.

Importantly, the definition of the cluster richness λ from the redMaPPer algorithm may differ from one dataset to another. Before comparing to our results, we implement this correction using the conversion obtained by McClintock et al. (2019):

$$\begin{aligned} \lambda_{\text{DES SV}} &= (1.08 \pm 0.16)\lambda_{\text{DES Y1}} \\ \lambda_{\text{SDSS}} &= (0.93 \pm 0.14)\lambda_{\text{DES Y1}} \end{aligned} \tag{4.11}$$

where the number presented as the uncertainty is actually the standard deviation in the richness ratio (thus, the uncertainty on the mean conversion factor is tiny in comparison). We have applied these corrections to bring all results to the space of our analysis.

Discussion of the mass trend parameter B_λ

Our mass trend shows good agreement with the results obtained by Saro et al. (2015), which is based on measurements of a cross-matched sample of SZE selected galaxy cluster candidates from the South Pole Telescope 2500 deg^2 SPT-SZ survey and the optically selected redMaPPer clusters from the Dark Energy Survey Science Verification (DES-SV) data. We also find good agreement with the scaling relation obtained by Baxter et al. (2016) and Baxter et al. (2018), where the first is based on cluster clustering using SDSS data, and the second on Cosmic Microwave Background (CMB) lensing measurements from SPT in combination with DES Y1 redMaPPer clusters.

On the other hand, our results are in disagreement with those of Simet et al. (2017), based on redMaPPer clusters found in the Sloan Digital Sky Survey, and of McClintock et al. (2019), obtained analyzing redMaPPer galaxy clusters identified in the Dark Energy Survey Year 1 data. While our analysis is performed on ensembles consisting of single clusters, these two analyses made use of stacked weak lensing data. In fact, neither of these analyses aimed to account for the Eddington bias and, therefore, they do not solve for the underlying richness-mass relation as we have done. Rather, they fit the mean λ within bins of lambda and redshift to the mean weak lensing mass associated with each bin. Because the Eddington bias is a function of the scatter in λ and the effective slope of the mass function at the corresponding mass, ignoring the Eddington bias correction will lead to systematic errors in the redshift and mass trends. We estimate that the Eddington bias correction will impact the mass and redshift trends with $\delta B_\lambda = +0.04$ and $\delta\gamma_\lambda = +0.09$, respectively, where δ is defined as the value of the parameter after applying the bias correction minus the one before the correction. With these corrections, the expected parameters for the underlying λ-mass relation would be $B_\lambda = 0.77$ and $\gamma_\lambda = -0.01$. These are still offset significantly from our measured values at $\Delta B_\lambda = -0.21 \pm 0.08$ (2.7σ) and $\Delta\gamma_\lambda = +1.12 \pm 0.60$ (1.9σ), and so clearly the Eddington bias is not large enough to explain the differences between the two results.

We note that redMaPPer optical selection and RASS X-ray selection followed by cross-matching to redMaPPer (i.e., the CODEX sample we analyze here) will not generally lead to similar levels of contamination by random superpositions. Moreover, contamination would be expected to have a different impact on a stacked weak lensing analysis than

on a cluster by cluster dynamical analysis like that carried out here. Thus, in principle, differences in the λ-mass relations constrained from these two different approaches can be used to shed light on the differences in contamination.

The contamination of optically selected cluster samples by projected collections of passive galaxies in low mass groups and isolated systems has long been a concern (Gladders et al., 2007; Song et al., 2012a; Costanzi et al., 2018), with estimates of contamination fractions reaching as high as $\sim$50%. Within X-ray imaging surveys like those employing pointed PSPC observations with $\sim 25''$ FWHM imaging (e.g. Vikhlinin et al., 1998; Clerc et al., 2018), the selection of X-ray sources exhibiting extended emission has been shown to deliver contamination at the $\sim$10% level. Within the lower quality RASS imaging, where there is generally no extent information for the faint CODEX sources, the contamination is driven by random superpositions between the faint X-ray sources ($\sim$ 90% are AGN or stars) and the ubiquitous red sequence optical candidate clusters identified by redMaPPer (see detailed discussion of this problem and the description of a method to control this contamination in Klein et al., 2018, Klein et al, in prep.).

Within a stacked weak lensing analysis, these contaminating low mass systems would likely suppress the mass at a given λ, and a mass dependent contamination that increases toward low λ, as suggested by some studies (Saro et al., 2015), could lead to a significant bias to low values in the mass slope B_λ. Within this context, it is interesting to note that the disagreement in the λ-mass relations between McClintock et al. (2019) and our our analysis is largest at low lambda.

For the CODEX sample, the random superpositions are not necessarily contaminants in a study of the λ–mass relation, because many of these random superpositions are of X-ray AGN projected to lie near true red sequence clusters on the sky. Subsequent spectroscopic followup of these systems, whether the X-ray emission is AGN or cluster dominated, leads to dynamical sampling of clusters and groups, with less impact from the tail of low mass, contaminating structure projections than in the case of the purely optically selected sample. Spectroscopic followup further reduces the contamination, because those systems that are loose projections can in many cases be separated out from the true, collapsed halos during the SPIDERS validation procedure (see also detailed spectroscopic studies of redMaPPer systems in Sohn et al., 2018; Rines et al., 2018).

Because our dynamical analysis uses (weak) mass information from all individual systems, the impact of the final remaining contamination in the CODEX calibration of the λ-mass relation, which would tend to be sampled with smaller numbers of spectroscopic redshifts, would then be further reduced. Thus, because both methods– optical cluster selection + stacked weak lensing and RASS+optical redMaPPer + dynamics– are subject to different systematic effects, we have a potential explanation for the different mass slopes observed in the two analyses. Further work using structure formation simulations or generation of realistic mocks including the appropriate contamination effects would be required to quantify these effects and understand the differences in detail. Supplementing this with dense spectroscopic studies of redMaPPer samples to better understand the nature of the projection and contamination issues will also be very helpful (Sohn et al., 2018; Rines et al., 2018).

Finally, we compare our scaling relation amplitudes and mass trends with those obtained by two recent low redshift ($z \leq 0.33$) SDSS based analyses. Murata et al. (2018) perform a richness-mass scaling relation calibration using a joint measurement of the abundance and stacked cluster weak lensing profiles within the context of the cosmological parameters preferred by Planck CMB anisotropy (Planck Collaboration et al., 2016d). They determine a scaling relation that reproduces both the cluster counts and the lensing profiles but only at very large richness scatter $\sigma_{\ln\lambda|M} = 0.46$. Scatter of this scale predicts a non-negligible contribution of low-mass haloes ($M_{200\mathrm{m}} \lesssim 10^{13} M_{\odot}$) in the SDSS redMaPPer sample. Their interpretation is that this contamination could be due to projection effects that preferentially impact the low richness portion of the sample ($20 \leq \lambda \leq 30$) or that the assumed Planck cosmology is different from the true underlying cosmology. We find good agreement with the mass trend of their results, but their amplitude is only about half of the value we find. The offsets in amplitude are not surprising given the very large differences in the scatter in the two analyses.

Jimeno et al. (2017) calibrate the mass-richness scaling relation using both the cluster correlation function and the cluster counts. They employ the N-body Millennium XXL simulations, updated to the Planck cosmology (Planck Collaboration et al., 2016d) to predict the distributions of clusters in richness. They first obtain two independent mass-richness relations using separately clustering and counts data, and afterwards perform a joint analysis. Interestingly, they find a 2.5σ tension between the amplitudes of the scaling relation in the two cases that weakens if they shift from the Planck cosmological parameters to those from the WMAP mission (Spergel et al., 2003). The joint constraints on the amplitude and mass trend of the mass-richness relation are in good agreement with our results.

Overall, the agreement with the counts+clustering analysis is encouraging, suggesting that their modeling of the redMaPPer selection and contamination cannot be far off. However, the counts+stacked weak lensing analysis seems to provide further indications that projection effects in the redMaPPer sample may be responsible for differences between stacked weak lensing constraints and measurements of the true underlying richness-mass relation from direct mass measurements (our analysis), from counts or from cluster clustering. Commonalities between the impact of correlated large scale structure on weak lensing and richness measurements may lie at the heart of these differences.

Discussion of the redshift trend parameter γ_λ

Our constraint on the redshift trend of the $\lambda - M_{200\mathrm{c}} - z$ relation shows a stronger negative trend $\gamma_\lambda = -1.13 \pm 0.33 \pm 0.49$ than found in previous analyses (Fig. 4.8), which have provided no significant evidence of a redshift trend (Saro et al., 2015; McClintock et al., 2019). The behavior we see in the CODEX sample would be expected if there were an increasing fraction of red sequence (RS) galaxies over cosmic time, with no evolution in the overall halo occupation number N_{200} of galaxies within the virial region above a particular stellar mass or luminosity cut. The redshift trend we measure is in rough agreement with results from Hennig et al. (2017), a study of the galaxy populations in 74 SPT clusters

whose redshifts extend to $z \sim 1.1$ and that were imaged as part of the DES SV survey. They find that the number of red sequence galaxies $N_{200,\mathrm{RS}}$ brighter than $m_* + 2$ and within r_{200c} decreases with redshift at fixed mass as $N_{200,\mathrm{RS}} \propto (1+z)^{-0.84\pm0.34}$, corresponding to an evolution of the red sequence fraction within r_{200c} going as $f_{\mathrm{RS}} \propto (1+z)^{-0.65\pm0.21}$. This evolution is less steep than the λ-mass evolution we observe here, but the two results are statistically consistent with a difference of 0.48 ± 0.63.

In contrast, the McClintock et al. (2019) and Saro et al. (2015) results show no redshift trend with $\gamma_\lambda = -0.22 \pm 0.22$ and 0.60 ± 0.63, respectively. These results differ from our measurement at 1.5σ (0.91 ± 0.63) and 2σ (1.73 ± 0.86), respectively. Interestingly, as discussed in Section 4.4.2, our measured redshift trend is closer to that measured in the other two analyses when we include more high redshift clusters that are sampled by smaller numbers of spectroscopic redshifts. Clearly, further study is needed to better understand whether there is a difference in the redshift trend inferred from dynamical masses and to pinpoint any underlying causes.

4.5 Conclusions

In this paper, we present a richness-mass-redshift scaling relation calibration using galaxy dynamical information from a sample of 428 CODEX galaxy clusters. These are X-ray selected systems from RASS that have red sequence selected redMaPPer optical counterparts within a search radius of $3'$. Our sample has redshifts up to $z \sim 0.66$ and optical richnesses $\lambda \geq 20$. The spectroscopic follow-up comes from the SPectroscopic IDentification of eRosita Sources (SPIDERS) survey, resulting in 7807 red member galaxies after interloper rejection and the exclusion of all systems with fewer than 10 member redshifts.

We study the λ-$M_{200c} - z$ relation by extracting the likelihood of consistency between the velocity sample for each individual cluster and the modeled projected phase space velocity distribution for a cluster of λ and redshift z inferred mass M_{200c}. The modeling is carried out using a Jeans analysis based on the code MAMPOSSt (Mamon et al., 2013), which allows us to build the projected phase space velocity distributions for a clusters of a particular mass, given a range of models for the orbital anisotropy of the galaxies. In our analysis, we adopt an NFW mass profile and employ five different velocity dispersion anisotropy profiles. Furthermore, we adopt an NFW profile for the red galaxy tracer population with concentration $c = 5.37$ (Hennig et al., 2017, and Section 4.2.5). We combine results from the different anisotropy models by performing Bayesian model averaging, allowing us to effectively marginalize over the orbital anisotropy of the spectroscopic galaxy population.

We model the scaling relation as $\lambda \propto \mathrm{A}_\lambda {M_{200c}}^{\mathrm{B}_\lambda}(1+z)^{\gamma_\lambda}$ (equation 4.5). As described in Section 4.4.1, we apply corrections for the Eddington bias and for the Malmquist bias. Results are presented in Table 4.2. For clusters at our pivot redshift of $z_{\mathrm{piv}} = 0.18$ and pivot mass of $M_{\mathrm{piv}} = 3 \times 10^{14} M_\odot$, we find our constraints on the scaling relation to be as

follows: the normalization A_λ, mass slope B_λ and redshift slope γ_λ are

$$\begin{aligned} A_\lambda =& 38.6^{+3.1}_{-4.1} \pm 3.9, \\ B_\lambda =& 0.99^{+0.06}_{-0.07} \pm 0.04, \\ \gamma_\lambda =& -1.13^{+0.32}_{-0.34} \pm 0.49. \end{aligned} \tag{4.12}$$

As discussed in Section 4.4.2, the quoted uncertainties include a 10% systematic uncertainty on the dynamical mass that is applied wholely to the scaling relation amplitude (see study of systematics in Mamon et al., 2013) and a systematic uncertainty of 0.49 on the redshift trend γ_λ, that arises from sensitivity in our redshift trend parameter to cuts on the cluster sample according to the number of member galaxies with spectroscopic redshifts.

Our results on the mass trend of the λ-mass scaling relation are in generally good agreement with previous studies of the mass dependence of the halo occupation number, or the number of cluster galaxies within a common portion of the luminosity function (often $m_* + 2$) and within a common portion of the cluster virial region (typically defined using r_{500} or r_{200}) (Lin et al., 2004; Hennig et al., 2017). This is an indication that the redMaPPer algorithm is effective at selecting cluster galaxies over a common portion of the virial region and that the galaxy red sequence fraction is not a strong function of cluster mass in this mass range.

Moreover, our results are in good agreement with those from previous studies of the λ-mass relation using SPT selected clusters that have been cross-matched with DES SV identified optical systems (Saro et al., 2015). We are also consistent with the value of the mass trend measured using cluster clustering in SDSS (Baxter et al., 2016) and CMB lensing of the DES Yr 1 redMaPPer sample using SPT (Baxter et al., 2018). On the other hand, our results are in disagreement with a study of redMaPPer clusters detected in SDSS data (Simet et al., 2017) and show a $\sim 2.7\sigma$ tension with the constraints obtained from redMaPPer galaxy clusters identified in the DES Y1 data (McClintock et al., 2019). Both of these latter results arose from the analysis of stacked weak lensing signatures, and neither analysis sought to obtain the true underlying λ-mass relation after correction for the Eddington bias. As discussed in Section 4.4.3, the Eddington bias correction would not be large enough to explain the difference. We suggest instead that the difference is reflective of the likely differences in the contamination of a pure RedMaPPer sample and our CODEX sample, which is first X-ray selected and then cross-matched to the RedMaPPer candidates within $3'$ radius.

In Section 4.4.3 we also discuss two scaling relation calibrations that adopt redMaPPer counts together with either stacked weak lensing or cluster clustering to calibration the richness-mass relation. Inferring cluster mass information from the counts such as in those two analyses requires an accurate description of the contamination or projection effects in the redMaPPer sample. Interestingly, our dynamical mass calibration results are in good agreement with the counts+clustering analysis (Jimeno et al., 2017), but not with the counts+stacked weak lensing analysis (Murata et al., 2018), where the authors find a dramatically larger scatter in richness-mass is required to bring their weak lensing and counts constraints on cluster masses into agreement.

The redshift trend γ_λ of our richness-mass relation shows a strong negative trend where λ at fixed mass decreases with redshift. This result can be interpreted as an indication of the increasing fraction of cluster red sequence galaxies over cosmic time. As presented in Section 4.4.3, our results are somewhat steeper than but statistically consistent with those from Hennig et al. (2017), where they studied SPT selected clusters and found that the fraction of red sequence galaxies decreases with redshift, from $\sim 80\%$ at $z \sim 0.1$ to $\sim 55\%$ at $z \sim 1$, following the form $f_{\rm RS} \propto (1+z)^{-0.65\pm0.21}$. However, our measurement is steeper than other results showing little or no redshift trend in the λ-mass relation (Saro et al., 2015; McClintock et al., 2019), but the differences are only significant at 1.5 and 2σ, respectively. Further study of the redshift trend of the λ–mass relation is clearly warranted.

In addition, we test the impact of interesting selection effects on our results in Section 4.4.2. We show that negative covariance between the scatter in X-ray luminosity and the scatter in optical richness for clusters at the levels measured in the CODEX sample has negligible impact on the λ-mass relation.

In summary, dynamical masses are a powerful tool to gain information on the link between the masses of galaxy clusters and readily obtainable observables– even in the limit of large cluster samples with small spectroscopic samples available for each cluster. Further work to perform a dynamical analysis on numerical simulations of structure formation will be crucial to being able to properly assess the true precision and robustness of the dynamical masses and anisotropy measurements we seek to extract from the data. A better understanding of the expected variation of the velocity anisotropy profile, of the distribution of interlopers after cleaning and of the impact of departures from equilibrium on our Jeans analysis will be broadly helpful. Our analysis demonstrates that there is promise in the analysis of small per-cluster spectroscopic samples of the sort that will be delivered by future spectroscopic surveys like DESI (Levi et al., 2013), 4MOST (de Jong et al., 2012) and Euclid (Laureijs et al., 2011).

Acknowledgments

RC would like to thank Gus Evrard, Arya Farahi, Tom McClintock and Steffen Hagstotz for helpful discussions. The Munich group acknowledges the support by the DFG Cluster of Excellence "Origin and Structure of the Universe", the Transregio program TR33 "The Dark Universe", the MPG faculty fellowship program and the Ludwig-Maximilians-Universität Munich. RC acknowledges participation in the IMPRS on Astrophysics at the Ludwig-Maximilians University and the associated financial support from the Max-Planck Society. RC and VS acknowledge support from the German Space Agency (DLR) through *Verbundforschung* project ID 50OR1603. AS is supported by the ERC-StG "ClustersXCosmo", grant agreement 716762. AB acknowledges the hospitality of the LMU and partial financial support from PRIN-INAF 2014 "Glittering kaleidoscopes in the sky: the multifaceted nature and role of Galaxy Clusters?", P.I.: Mario Nonino.
Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.
SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Chapter 5

Mass Calibration of the CODEX Cluster Sample using SPIDERS Spectroscopy - II. The X-ray Luminosity-Mass Relation

R. Capasso[1,2], J. J. Mohr[1,2,3], A. Saro[1,2,4], A. Biviano[4], N. Clerc[3]2[,5], A. Finoguenov[3,6]2, M. Klein[1,3], S. Grandis[1,2]

2019, to be submitted to Monthly Notices of the Royal Astronomical Society
(This draft is currently being reviewed by the collaborators)

[1]Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany
[2]Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany
[3]Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 85748 Garching, Germany
[4]INAF-Osservatorio Astronomico di Trieste via G.B. Tiepolo 11, 34143 Trieste, Italy
[5]IRAP, Université de Toulouse, CNRS, UPS, CNES, Toulouse, France
[6]Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2a, FI-00014 Helsinki, Finland

ABSTRACT

We perform the calibration of the X-ray luminosity–mass scaling relation on a sample of 344 CODEX clusters with $z < 0.66$ using the dynamics of their member galaxies. Spectroscopic follow-up measurements have been obtained from the SPIDERS survey, leading to a sample of 6,658 red member galaxies. We use the Jeans equation to calculate halo masses, assuming an NFW mass profile and analyzing a broad range of anisotropy profiles. With a scaling relation of the form $L_X \propto A_X M_{200c}^{B_X} E(z)^2(1+z)^{\gamma_X}$, we find best fit parameters $A_X = 5.7^{+0.4}_{-0.5}(\pm 0.6) \times 10^{43}\,\mathrm{erg\,s^{-1}}$, $B_X = 2.5 \pm 0.2(\pm 0.06)$,$\gamma_X = -2.6^{+1.1}_{-1.2}(\pm 0.76)$, where we include systematic uncertainties in parentheses and for a pivot mass and redshift of $3 \times 10^{14} M_\odot$ and 0.16, respectively. We compare our constraints with previous results, and we combine our sample with the SPT SZE–selected cluster subsample observed with XMM-*Newton* to improve constraints on the redshift evolution.

5.1 Introduction

Accurate mass estimates of galaxy clusters are of fundamental importance for both cosmological and astrophysical studies. Observational knowledge of the mass distribution of the dark and baryonic matter in clusters provides insights into their formation and evolution (see, e.g. Gao et al., 2004; Springel et al., 2001). On the other hand, number counts of galaxy clusters, sensitive to the amplitude of matter fluctuations, can provide constraints on various cosmological parameters (e.g., White et al., 1993b; Haiman et al., 2001; Mantz et al., 2015; Bocquet et al., 2018). Studies of the link between the observable features of haloes and the underlying matter distribution are thus essential.

An efficient use of clusters as cosmological probes requires a low-scatter mass proxy (Lima & Hu, 2005; Allen et al., 2011). A sample of galaxy clusters is generally affected by a number of biases, depending on the intrinsic covariance of the cluster observables, measurement uncertainties and selection effects (e.g. Pacaud et al., 2007; Mantz et al., 2010; de Haan et al., 2016). The combination of limited surveyed volume and source selection thresholds produce the well known Malmquist bias (Malmquist, 1920), truncating the scattered distributions of sources in the space of observables. As a consequence, luminosity or flux selected samples are typically biased towards low masses where the selection is returning only a fraction of the underlying cluster sample. This effect is enhanced by the so-called Eddington bias (Eddington, 1913). Because the number density of halos is a steeply falling function of their mass (e.g. Tinker et al., 2008; Bocquet et al., 2016), the presence of scatter in the relationship between the selection observable (i.e., flux or luminosity) and mass will cause low-mass clusters to preferentially up-scatter, leading to a bias in the mass associated with the observable (Mortonson et al., 2011). An accurate calibration of cluster scaling relations requires control over these biases.

Many different mass proxies have been used over the years, including thermal Sunyaev–Zel'dovich effect (SZE) measurements (Staniszewski et al., 2009; Planck Collaboration et al., 2014b; Hasselfield et al., 2013), weak gravitational lensing features (Corless & King, 2009; Becker & Kravtsov, 2011; Dietrich et al., 2019), cluster velocity dispersions (Biviano

et al., 2006; Saro et al., 2013; Capasso et al., 2019), and X-ray luminosity and temperature (Vikhlinin et al., 2009; Mantz et al., 2010). A combination of multiple, independent mass proxies help mitigate systematic errors (Bocquet et al., 2015; McClintock et al., 2019; Baxter et al., 2018; Farahi et al., 2018; Bocquet et al., 2018). In a companion paper (Capasso et al., 2018, hereinafter C18) we performed the dynamical mass calibration exploiting the optical richness of a sample of 428 CODEX (COnstrain Dark Energy with X-ray clusters; Finoguenov, in prep) clusters, constraining the amplitude of the λ–mass relation with a $\sim$12% accuracy.

Following C18, we calibrate the X-ray luminosity–mass–redshift scaling relation by exploiting the information residing in the observed projected phase space (distribution in line of sight velocities and projected radius) of the cluster member galaxies. We use a modification of the MAMPOSSt technique (Modeling Anisotropy and Mass Profiles of Observed Spherical Systems; Mamon et al., 2013), based on the Jeans equation (Binney & Tremaine, 1987), to simultaneously determine the dynamical cluster masses and the parameters of the scaling relation. The MAMPOSSt code has been successfully used to investigate the internal dynamics of clusters, determining their masses and velocity anisotropy profiles (e.g. Biviano et al., 2013, 2017; Munari et al., 2014; Capasso et al., 2019).

We perform this analysis on the CODEX cluster catalog, which consists of ROSAT All-Sky Survey (RASS) X-ray cluster candidates having optical counterparts in SDSS imaging data identified using the RedMaPPer algorithm (the red-sequence Matched-filter Probabilistic Percolation algorithm, Rykoff et al., 2014). A subset of this sample has been spectroscopically studied within the SPectroscopic IDentification of eRosita Sources (SPIDERS) survey (Clerc et al., 2016). The analysis we carry out focuses on a sample of 344 CODEX clusters with a corresponding sample of $\sim$6600 red member galaxies with measured redshifts. The clusters span the redshift range $0.03 \leq z_{\rm c} \leq 0.66$, with richnesses $20 \leq \lambda \leq 230$ and X-ray luminosities $4.5 \times 10^{42} \leq L_{\rm X}/({\rm erg\,s^{-1}}) \leq 3.2 \times 10^{45}$.

The paper is organized as follows. In Section 5.2 we detail the theoretical framework. In Section 5.3 we present the data set used in this analysis and the selection criteria. The likelihood model used to constrain the $L_{\rm X}$–mass–redshift scaling relation is described in Section 5.4, followed by the outcome of our calibration, and a discussion of a range of systematic uncertainties. We present our conclusions in Section 5.5.

Throughout this paper we assume a flat ΛCDM cosmology with a Hubble constant $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and a matter density parameter $\Omega_{\rm M} = 0.3$. Cluster masses ($M_{\rm 200c}$) are defined within $r_{\rm 200c}$, the radius of the sphere inside which the cluster overdensity is 200 times the critical density of the Universe at the cluster redshift. We refer to $r_{\rm 200c}$ as the virial radius. All quoted uncertainties are equivalent to Gaussian 1σ confidence regions, unless otherwise stated.

5.2 Theoretical Framework

We use dynamical constraints on a large ensemble of clusters to constrain the underlying halo masses, thereby enabling measurement of the luminosity–mass-redshift relation. To do this, we perform a dynamical analysis based on the application of the Jeans equation to spherical systems (Binney & Tremaine, 1987). The Jeans equation allows us to define the mass distribution $M(r)$ of a cluster as

$$\frac{GM(<r)}{r} = -\sigma_r^2 \left(\frac{d\ln\nu}{d\ln r} + \frac{d\ln\sigma_r^2}{d\ln r} + 2\beta \right), \tag{5.1}$$

with $\nu(r)$ being the number density profile of the tracer galaxy population, $\sigma_r(r)$ the radially dependent component of the velocity dispersion along the spherical coordinate r, $M(<r)$ the enclosed mass within radius r, G Newton's constant, $\beta(r) \equiv 1 - (\sigma_\theta^2/\sigma_r^2)$ the radially dependent velocity dispersion anisotropy, and σ_θ the tangential component of the velocity dispersion.

Equation 5.1 can thus be used to estimate the mass distribution of a spherical system. However, the only observables we can directly obtain are projected quantities: the surface density profile of the galaxy distribution, the rest-frame LOS velocities and the radial separation of each galaxy from the cluster center. Given the typical observational uncertainties on the line of sight velocity distribution, it is not possible to uniquely derive the mass distribution of a galaxy cluster (Merritt, 1987), leading to a degeneracy between the mass and the velocity anisotropy profiles.

In this work, we address this problem by implementing the Modeling Anisotropy and Mass Profiles of Observed Spherical Systems algorithm (hereafter MAMPOSSt; for full details please refer to Mamon et al., 2013). This method consists in determining the mass and anisotropy profiles of a cluster in parametrized form by performing a likelihood exploration of the distribution of the cluster galaxies in projected phase space, comparing it to the theoretical distribution predicted from the Jeans equation for these models. This method thus requires adopting parametrized models for the number density, mass, and velocity anisotropy profiles $\nu(r)$, $M(r)$, $\beta(r)$.

As addressed in Section 5.3.6, because our spectroscopic dataset is likely to suffer from radially dependent incompleteness, we adopt the number density profile derived from a study of red sequence galaxies in SZE selected clusters (Hennig et al., 2017).

Regarding our choice of the mass and velocity anisotropy profiles, we follow our previous work C18. We refer to that study for a more detailed description. In the next section we summarize the main features.

5.2.1 Mass and anisotropy profiles

Driven by both numerical studies of structure formation and observational results, we adopt the mass model introduced by Navarro et al. (1996, NFW), which is fully described by two parameters: the virial radius r_{200}, and the scale radius r_s, which is the radius at which the logarithmic slope of the density profile is -2. Numerous observational studies

have indeed found the mass distributions of clusters to be well described by this model (Carlberg et al., 1997b; van der Marel et al., 2000b; Biviano & Girardi, 2003; Katgert et al., 2004b; Umetsu et al., 2011).

On the other hand, due to the lack of published studies providing strong predictions for the radial form of the velocity anisotropy profile $\beta(r)$, we consider five models that have been used in previous MAMPOSSt analyses, described also in Capasso et al. (2019): (1) constant anisotropy model (C), (2) Tiret anisotropy profile (Tiret et al., 2007, T), (3) Mamon & Łokas (2005) profile (MŁ), (4) Osipkov-Merritt anisotropy profile (Osipkov, 1979; Merritt, 1985, OM), and (5) a model with anisotropy of opposite sign at the center and at large radii (O).

Therefore, given a mass for each cluster we predict the projected phase space distribution of the observed dynamical dataset by running MAMPOSSt with 3 free parameters: the virial radius r_{200c}, the scale radius r_s of the mass distribution, and a velocity anisotropy parameter θ_β. The latter represents the usual $\beta = 1 - (\sigma_\theta^2/\sigma_r^2)$ for the first three models (C, T, O), while for the MŁ and OM models it defines a characteristic radius $\theta_\beta = r_\beta$.

5.2.2 Bayesian model averaging

As described above, we employ five velocity anisotropy models when estimating the projected phase space distribution of member galaxies for each cluster. Because we cannot strongly reject any of the models, we combine the results obtained from each anisotropy model $\beta(r)$ by merging their constraints, exploiting the Bayesian model averaging technique (see C18, Capasso et al., 2019, for more details). In a nutshell, this method consists in assigning a weight to each model, according to how well the model fits the data. This weight is represented by the so-called Bayes factor (see Hoeting et al., 1999, and references therein).

Considering the 5 anisotropy models M_1, ..., M_5, the Bayes factor B_j of each model j is defined as the marginalized likelihood of the model $\mathcal{L}(D\,|M_j)$, also known as evidence, normalized by the likelihood of the most probable model. Specifically,

$$B_j = \frac{\mathcal{L}(D\,|M_j)}{\mathcal{L}(D\,|M_{\rm max})}, \tag{5.2}$$

where $M_{\rm max}$ indicates the model with the highest marginalized likelihood, $\mathcal{L}(D\,|M_j) = \int \mathcal{L}(D\,|\theta_j, M_j) P(\theta_j\,|M_j)\, d\theta_j$, $\mathcal{L}(D\,|\theta_j, M_j)$ is the likelihood of the data D given the model parameters θ_j, and $P(\theta_j\,|M_j)$ is the prior.

The average posterior distribution of the fitted scaling relation parameters is then given by the weighted average of the posterior distributions of each model, with the Bayes factor as weight. This Bayesian model averaging is performed by means of the multimodal nested sampling algorithm MultiNest (Feroz & Hobson, 2008; Feroz et al., 2009, 2013), providing us with the evidence for each model.

5.3 Data

We perform our analysis on a subset of CODEX galaxy clusters observed within the SPIDERS survey (Clerc et al., 2016), which provides us with the spectroscopic galaxy sample. The CODEX sample is based on ROSAT All-Sky Survey (RASS, see Voges et al., 1999) selected clusters, cross–matched with nearby optically selected systems identified using the redMaPPer (the red-sequence Matched-filter Probabilistic Percolation, Rykoff et al., 2014) algorithm applied to the Sloan Digital Sky Survey IV (SDSS-IV, see Dawson et al., 2016; Blanton et al., 2017) optical imaging data. A full description of the dataset construction and features are described in C18. In the following section we summarize the main elements of the dataset.

5.3.1 The CODEX sample

The CODEX cluster sample combines ROSAT X-ray cluster candidates with optical selected cluster candidates identified using redMaPPer. First of all, RASS data are searched for all X-ray sources with detection significance S/N$>$ 4. Then, redMaPPer is run on the SDSS imaging data around each of these sources. RedMaPPer is an optical cluster-finding algorithm based on the red sequence technique, built around the richness estimator of Rykoff et al. (2012). This step thus allows the identification of candidate clusters with a red-sequence, constituting a collection of passive galaxies at a common redshift. The redMaPPer algorithm provides an estimate for the cluster photometric redshift, an estimation of the optical richness and an optical cluster center. In cases of multiple optical counterparts meeting these criteria, the counterpart having the highest richness is assigned to the RASS X-ray source. The updated optical cluster position allows the identification of a revised red-sequence, providing the final estimate of the cluster photometric redshift and richness. Finally, RASS count-rates provide an estimate for the X-ray properties of the clusters. Assuming a model for the X-ray spectral emissivity, imposing a minimal S/N threshold of 1.6 to have optimized apertures, we calculate the aperture-corrected cluster flux f_X and [0.1-2.4] keV luminosities L_X.

The final CODEX sample is then characterized by X-ray detected clusters, with estimated redshift, optical richness, optical cluster center, and X-ray luminosity. Follow-up observations obtained with the SPIDERS survey, described below, finally provide us with spectroscopic redshift redshift measurements of cluster member galaxies.

5.3.2 The SPIDERS spectroscopic sample

The SPIDERS survey is designed to obtain homogeneous and complete spectroscopic follow-up of X-ray extragalactic sources lying within the SDSS-IV imaging footprint, with the aim of confirming galaxy cluster candidates and of assigning a precise redshift measurement. In particular, this survey was conceived to obtain follow-up observations of X-ray extended sources extracted from the all sky X-ray eROSITA survey (extended ROentgen

Survey with an Imaging Telescope Array; Predehl et al., 2010; Merloni et al., 2012). However, prior to the launch of eROSITA, the bulk of the SPIDERS program galaxy clusters will be those identified in the shallower RASS and sparser XMM-Newton data. At the time this paper is being written, the observations of these clusters have already been completed. No further galaxy spectroscopic redshifts will be assigned to them during the final stages of the SDSS-IV program.

The target selection is perform so as to optimize the number of spectroscopically confirmed clusters. As a first step, the redMaPPer membership probability is used to assign priorities to potential targets, ranking galaxies within each cluster. The pool of targets along with the priority flag is then submitted to the eBOSS tiling algorithm. The eBOSS spectroscopic pipeline is then employed to produce the final data reduction and spectral classification.

For each cluster, an automatic procedure assigns membership of red-sequence galaxies with measured redshifts. This is performed through an iterative clipping procedure. Members with velocity rest-frame velocities (relative to the first guess cluster redshift) greater than 5000 km/s are rejected. The remaining potential members are used to estimate the velocity dispersion of the cluster. A 3σ clipping is then applied, rejecting objects lying further away than 3 times the velocity dispersion from the mean velocity.

In the course of this iterative procedure, a few problematic cases typically occur. For example, fewer than 3 members are sometimes assigned to a cluster, and sometimes the initial 5000 km/s clipping rejects all members. In such cases the problematic cluster is flagged and visually inspected by independent inspectors. This final validation may lead to the inclusion or removal of members, as well as the identification of other structures lying along line-of-sight of the cluster. Final cluster redshift estimates are based on the bi-weight average (Beers et al., 1990) of all galaxies selected as cluster members, if at least 3 members are assigned to the cluster. The cluster redshift statistical uncertainty is typically $\Delta_z/(1+z) \lesssim 10^{-3}$.

Finally, the updated cluster spectroscopic redshifts are used to update the measurement of X-ray cluster properties. Assuming the standard flat ΛCDM cosmological model (Hubble constant $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, and matter density parameter $\Omega_\mathrm{M} = 0.3$, ROSAT fluxes are converted into rest-frame [0.1-2.4] keV luminosities. The typical measurement uncertainty on the luminosities is $\approx 35\%$, as computed from the Poissonian fluctuation in the associated ROSAT X-ray photons (see Mirkazemi et al., 2015).

5.3.3 Final spectroscopic cluster member sample

Before proceeding with our analysis, we apply some additional cuts to the SPIDERS spectroscopic sample. First of all, to avoid merging systems, we only use clusters which do not have any other component along the line of sight. As we are carrying out a Jeans analysis, based on the assumption of dynamical equilibrium, we restrict our analysis to the cluster virial region ($R \leq r_{200c}$). Moreover, we exclude the very central cluster region ($R \leq 50$kpc), to avoid the inclusion of the central BCG, in which merger and dissipation processes could be ongoing, and to account for the positional uncertainties of cluster centers. Our final

spectroscopic dataset consists of 705 galaxy clusters, for a total of $\approx 11,400$ candidate cluster members, with a median redshift $z = 0.21$ and spanning an X-ray luminosity range $4.5 \times 10^{42} \leq L_{\rm X}/({\rm erg\,s^{-1}}) \leq 3.2 \times 10^{45}$, and a richness range $20 \leq \lambda \leq 230$.

We take the SPIDERS validated redshifts and redMaPPer positions to calculate the observable needed for our analysis: the galaxy projected clustercentric distance R and the rest-frame line of sight (LOS) velocity $\rm v_{rf}$. We convert galaxies and clusters redshifts, $z_{\rm gal}$ and $z_{\rm c}$, into velocities, $\rm v(z_{gal})$, and $\rm v(z_c$. Rest-frame velocities $\rm v_{rf}$ are then extracted from these quantities as $\rm v_{rf} \equiv (v(z_{gal}) - v(z_c))/(1 + z_c)$.

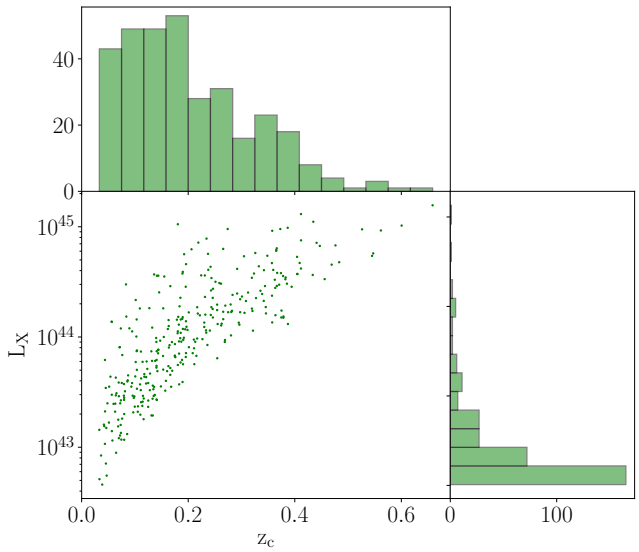

Figure 5.1: Distribution of X-ray luminosity $L_{\rm X}$ and cluster redshift $z_{\rm c}$ of the final cluster sample.

5.3.4 Interloper rejection

As described in Section 5.3.2, the SPIDERS automated procedure assesses membership for each galaxy in each cluster. However, interloper galaxies could still be present. These are galaxies that in projection are inside the cluster virial region, but do not actually lie inside it. We identify these objects by means of the "Clean" method (Mamon et al., 2013), based on the comparison between the location of the galaxies in the projected phase space and the expected maximal line of sight velocity at each projected radius. As we do not have enough spectroscopic redshifts to perform this method accurately for each individual cluster, we divide our sample into 15 equally spaced λ bins, building a composite cluster in each bin. The composite clusters are built by stacking in metric radius [Mpc], without applying any scaling in velocity. We then perform the interloper rejection in each of them separately.

The cleaning is performed in several steps. For each composite cluster, the LOS velocity dispersion $\sigma_{\rm LOS}$ is used to estimate the cluster mass $M(r)$, using a scaling relation calibrated using numerical (e.g., Saro et al., 2013), and assuming an NFW mass profile with

concentration sampled from the mass–concentration relation. Then, assuming the MŁ velocity anisotropy profile model, and given the cluster $M(r)$, an LOS velocity dispersion profile $\sigma_{\rm LOS}(R)$ is calculated. Finally, galaxies with $|{\rm v_{rf}}| > 2.7\sigma_{\rm LOS}$ at any clustercentric distance are iteratively rejected (see Mamon et al., 2010a, 2013).

After the removal of interlopers, our spectroscopic sample consists of 703 clusters and 9,121 red galaxies. We apply a further cut on this dataset: we only keep systems that have at least 10 spectroscopic members, $N_{\rm mem} \geq 10$. This decision is driven by our concern that good constraints on the cluster masses and scaling relation parameters could not be obtained from clusters having very small numbers of spectroscopic members. We explore the impact of this cut in Section 5.4.2. After this cut, we are left with 428 clusters and 7807 red galaxies, with a median redshift, richness, and luminosity of $z = 0.16$, λ=41, and ${\rm L_X} = 9.2 \times 10^{43} {\rm erg\,s^{-1}}$, respectively. Fig. 5.1 shows the distributions of cluster redshift and X-ray luminosity of the final sample.

We note that, even after this cleaning procedure, there is still a degree of contamination by interlopers. In general, galaxies lying outside the virial radius tend to have smaller peculiar velocities than those inside R_{200}. Galaxies close to the cluster turn-around radius will have negligible peculiar velocities, and will not be identified as interlopers by the method adopted here. An analysis of cosmological N-body simulations carried out by Saro et al. (2013) shows that, when passive galaxies are selected, this contamination is characteristically ∼20% for massive clusters ($M_{\rm 200c} \geq 10^{14} M_\odot$), increasing with decreasing cluster mass. Another analysis carried out by Mamon et al. (2010a) on hydrodynamical cosmological simulations shows that the distribution of interlopers in projected phase space is nearly universal, presenting only small trends with cluster mass. They find that, even after applying the iterative $2.7\sigma_{\rm LOS}$ velocity cut, the fraction of interlopers is still $23 \pm 1\%$ of all DM particles with projected radii within the virial radius, and over 60% between 0.8 and 1 virial radius.

5.3.5 Removing CODEX catalog contamination

When cross-matching X-ray selected candidates from RASS with optical systems from redMaPPer or other similar techniques, one must be careful to account for the contamination of the resulting cluster catalog by random superpositions of physically unassociated X-ray and optical systems along the line of sight (Klein et al., 2018). For the RASS imaging, where there is generally no extent information for the faint CODEX sources, the contamination is driven by random superpositions between the faint X-ray sources (∼ 90% are AGN or stars) and the ubiquitous red sequence optical candidate clusters identified by redMaPPer.

To exclude chance superpositions, we employ the method described in Klein et al. (2019). This decontamination method consists of evaluating, for a cluster candidate at redshift z and richness λ, the probability distribution of richness at that redshift for detected X-ray sources and that along random line of sights. Namely, we use the estimator $f_{\rm cont}$, which is defined as the ratio of the integral over the two distributions, above the observed λ of the candidate (see Fig. 6 and Eq. 10 in Klein et al., 2019). In particular, we

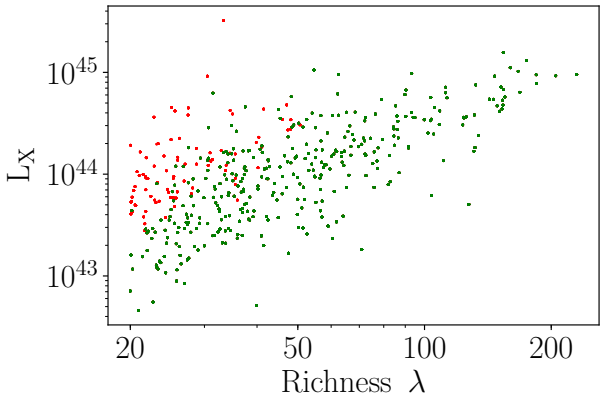

Figure 5.2: Distribution of X-ray luminosity and richness for our cluster sample. Red dots represent contaminated sources, which we exclude from our analysis. We imposed the cut $f_{\mathrm{cont,m}} < 0.05$, resulting in a catalogue with a 5% contamination fraction.

adopt the value of $f_{\mathrm{cont,m}}$, which uses the distribution of observed richness together with the weighted mean of random richness distributions, based on different count rates. This value is marginalized over the count rate in a particular redshift bin.

We perform a cut at $f_{\mathrm{cont,m}} < 0.05$, producing a sample with a 5% contamination fraction, which is independent of redshift. After the cut, our final sample consists of 344 galaxy clusters with a total of 6,658 cluster members, characterized by a median redshift $z = 0.16$ and a median X-ray luminosity $L_{\mathrm{X}} = 9 \times 10^{43}\,\mathrm{erg\,s^{-1}}$. Figure 5.2 shows the distribution of our sample as a function of richness and X-ray luminosity, highlighting in red clusters identified as contaminated. We exclude these objects from our analysis.

5.3.6 Galaxy number density profile

As showed in Section 5.2, the Jeans analysis requires knowledge of the 3D number density profile $\nu(r)$ of the tracer population, i.e. the red sequence member galaxies. The absolute normalization of the galaxy number density profile has no impact on our analysis, because only the logarithmic derivative of $\nu(r)$ enters the Jeans equation (see equation 5.1). On the other hand, a radially dependent incompleteness in the velocity sample would lead to a modification of the shape of the $\nu(r)$ profile, which would have an impact on our results. As the spectroscopic followup within SPIDERS will lead to a radially dependent incompleteness. we cannot simply adopt the spectroscopic sample to measure the number density profile of the tracer population. We therefore rely on a study of the galaxy populations in 74 SZE selected clusters from the SPT-SZ survey, imaged as part of the Dark Energy Survey Science Verification phase (Hennig et al., 2017). This study shows that the number density profile of the red sequence population is well fit by a Navarro, Frenk and White (NFW) model (Navarro et al., 1996) out to radii of $4r_{200c}$, with a concentration for cluster galaxies of $c_{\mathrm{gal}} = 5.37^{+0.27}_{-0.24}$. No statistically significant redshift or mass trends were identified in the radial distribution of red sequence galaxies for $z > 0.25$

A_{L_X}	B_{L_X}	γ_{L_X}	r_β	$\sigma_{\ln L_X^{int}}$
$\mathcal{U}(0.1, 2)$	$\mathcal{U}(1, 5)$	$\mathcal{U}(-7, 2)$	$\mathcal{U}(0.01, 10)$	$\mathcal{N}(0.27, 0.1^2)$

Table 5.1: Priors assumed for our analysis. $\mathcal{U}(i, j)$ refers to a uniform flat prior in the interval (i, j), while $\mathcal{N}(\mu, \sigma^2)$ indicates a Gaussian distribution with mean μ and variance σ^2.

and $M_{200c} > 4 \times 10^{14} M_\odot$. Therefore, we adopt the number density profile described by an NFW profile with the above-mentioned value of c_{gal} and a scale radius $r_{s,\,gal} = R_{200c}/c_{gal}$. Implicit in this approach is the assumption that the dynamical properties of our spectroscopic sample are consistent with those of the red sequence galaxy population analyzed by Hennig et al. (2017).

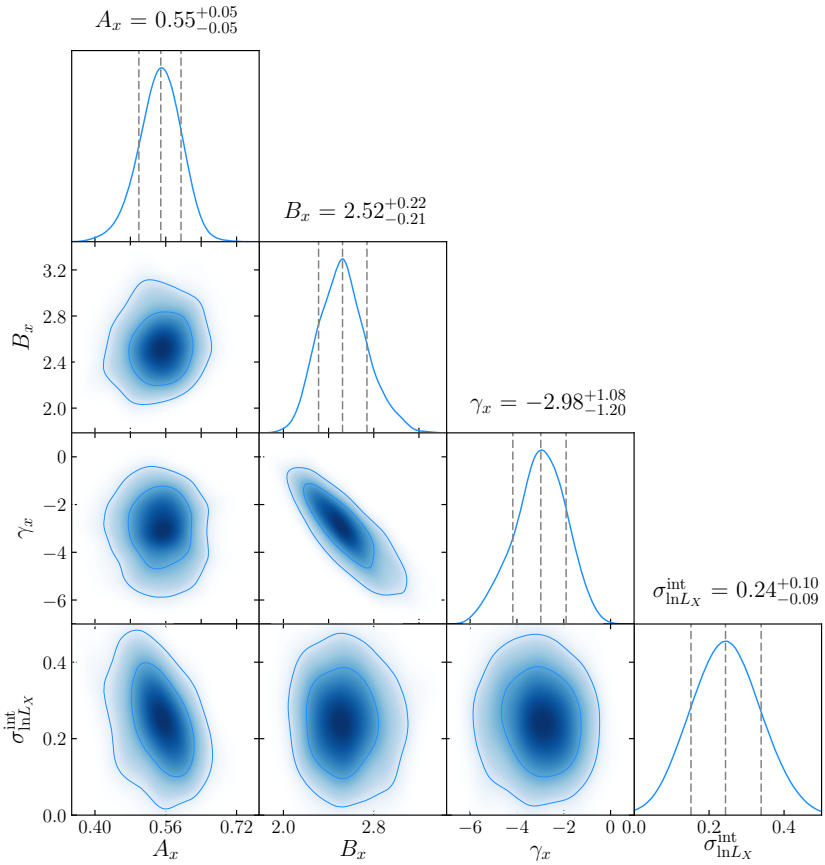

Figure 5.3: Posterior distribution of the scaling relation parameters. Contours show the 1σ, 2σ, and 3σ confidence regions.

Dynamical analysis using SPIDERS data	A_X	B_X	γ_X	C_X	$\sigma^{\rm int}_{\ln L_X}$
Baseline analysis	$0.57^{+0.04}_{-0.05} \pm 0.06$	$2.47^{+0.23}_{-0.19} \pm 0.08$	$-2.57^{+1.06}_{-1.17} \pm 0.76$	2	$0.24^{+0.10}_{-0.09}$
Combined analysis	$0.54^{+0.04}_{-0.05}$	$2.11^{+0.10}_{-0.10}$	$-0.52^{+0.39}_{-0.39}$	2	$0.24^{+0.09}_{-0.09}$
Previously published results	$L_X(3 \times 10^{14} M_\odot, 0.16)$	$M^{B_X}_{200c}$	$(1+z)^{\gamma_X}$	$E(z)^{C_X}$	
SPT + XMM-Newton (Bulbul et al., 2019)	0.58 ± 0.09	1.92 ± 0.18	0.004 ± 0.50	2	0.27 ± 0.10
WL + RASS (Nagarajan et al., 2018)	0.42 ± 0.27	1.62 ± 0.30	–	2	$0.75^{+0.19}_{-0.16}$
Chandra (Giles et al., 2017)	0.27 ± 0.13	1.96 ± 0.24	–	2	0.68 ± 0.11
Chandra + ROSAT (Mantz et al., 2016)	0.80 ± 0.35	1.35 ± 0.06	-0.65 ± 0.38	2.31 ± 0.06	0.42 ± 0.03
Chandra + ROSAT (Vikhlinin et al., 2009)	0.68 ± 0.21	1.63 ± 0.15	–	1.85 ± 0.42	0.396 ± 0.039

Table 5.2: X-ray luminosity-mass-redshift scaling relation parameters and intrinsic scatter from this analysis and the literature. Parameters are as defined in equation (5.3), and include the Eddington and Malmquist biases. Results from this analysis are showed with statistical uncertainties, together with a systematic mass uncertainties. In the comparison to previous results, the amplitude A_X column contains the luminosity at $M_{200c} = 3 \times 10^{14} M_\odot$ and $z = 0.16$. Conversions have been made to M_{200c} and from $E(z)$ to $(1+z)$ where needed. Note also that each of these studies was performed on a different range of mass and redshift.

5.4 Results

This section is dedicated to the results of the dynamical analysis. In the first subsection we present the method used to calibrate the $L_{\rm X}$-mass relation, and the results we obtain. We end the section with a comparison of our findings to those from previous studies, and we discuss the impact of the choice of the priors on our results.

5.4.1 Fitting Procedure

We model the relation between the X-ray luminosity, mass and redshift as

$$\frac{L_{\rm X}}{(10^{44}{\rm erg\,s^{-1}})} = {\rm A_X}\left(\frac{M_{\rm 200c}}{M_{\rm piv}}\right)^{\rm B_X}\left(\frac{E(z)}{E(z_{\rm piv})}\right)^{2}\left(\frac{1+z}{1+z_{\rm piv}}\right)^{\gamma_{\rm X}}, \tag{5.3}$$

where $\rm A_X$, $\rm B_X$, and $\gamma_{\rm X}$ are the amplitude, the mass slope and the redshift evolution slope. In this formulation, the redshift trend is expressed as both a function of z, and of the Hubble parameter $H(z) = H_0E(z)$, where, in a flat ΛCDM Universe, $E^2(z) = \Omega_{\rm m}(1+z)^3 + \Omega_\Lambda$ at late times. Therefore, we explicitly parametrize the cosmological dependence of the redshift evolution, while modeling departures from the self-similar evolution with a function $(1+z)^{\gamma_{\rm X}}$. Similar forms have been previously adopted to study the redshift and mass trends of the $L_{\rm X}$-mass relation (e.g., Bulbul et al., 2019). The pivot redshift is set to the be the median redshift of our sample, $z_{\rm piv} = 0.16$, while the mass pivot has been chosen in a way that minimizes the false degeneracy between the amplitude and the mass trend parameters, $M_{\rm piv} = 3\times10^{14}M_\odot$. We adopt a log-normal intrinsic scatter in $L_{\rm X}$ at fixed mass, $\sigma^{\rm int}_{\ln L_{\rm X}}$.

We follow the fitting framework presented in C18. Given the set of parameters $\boldsymbol{p}$, containing the 4 scaling relation parameters ($A_{\rm X}, B_{\rm X}, \gamma_{\rm X}, \sigma^{\rm int}_{\ln\lambda}$) and the anisotropy model parameter r_β, we calculate an initial mass $M_{\rm 200c,obs}$ using the scaling relation presented in Eq. 5.3. We then use the method of Mortonson et al. (2011) to estimate the Eddington bias correction caused by the interplay of the cluster mass function and scatter of the scaling relation. We assume the variance on the log-normal mass-observable relation to be $\sigma^2_{\ln M} = (1/B_{\rm X}\cdot\sigma_{\ln L_{\rm X}})^2$, where

$$\sigma^2_{\ln L_{\rm X}} = \left(\frac{\Delta L_{\rm X}}{L_{\rm X}}\right)^2 + {\sigma^{\rm int}_{\ln L_{\rm X}}}^2, \tag{5.4}$$

with $\Delta L_{\rm X}$ being the $L_{\rm X}$ measurement uncertainty divided by the observed luminosity. Assuming that the variance $\sigma^2_{\ln M}$ is small compared with the scale over which the local slope Γ of the mass function changes, the posterior mass distribution is a log-normal of the same variance $\sigma^2_{\ln M}$ with the mean shifting as $\ln < M_{\rm 200c,true} >= \ln < M_{\rm 200c,obs} > +\Gamma\sigma^2_{\ln M}$. We adopt this mass as input in MAMPOSSt, evaluating for each cluster the likelihood distribution in projected phase space. We combine the likelihoods calculated for each member galaxy in that cluster, such that the $i-$th term in the likelihood $\mathcal{L}_i$ contains the probability of observing the $i-$th cluster at redshift z_i, with mass $M_{\rm 200c,true,\,i}$ and X-ray luminosity $L_{\rm X,i}$, and the phase space of its member galaxies (clustercentric radii R^j and

rest-frame velocities v_{rf}^{j} of each $j-$th galaxy), given the scaling relation parameters A_X, B_X and γ_X, the anisotropy parameter r_β, and the intrinsic scatter $\sigma^{\rm int}_{\ln L_X}$:

$$\mathcal{L}_i = \prod_{j \in gal} \mathcal{L}(R^j, v_{rf}^j, L_{X,i}, z_i \mid \boldsymbol{p}). \tag{5.5}$$

The maximum likelihood solutions are obtained using the NEWUOA software (Powell, 2006). Priors on the parameters are assumed as follows (see Table 5.1): flat for the scaling relation parameters and for the anisotropy parameter, gaussian for the intrinsic scatter (mean $\mu = 0.27$ and variance $\sigma^2 = 0.1$, from Bulbul et al. (2019)).

The final likelihood for the total sample, for each set of scaling relation parameters $\boldsymbol{p}$, will then be obtained by combining the likelihoods for all the single clusters:

$$\mathcal{L} = \prod_{i \in clus} \mathcal{L}_i. \tag{5.6}$$

This procedure is carried on separately for each anisotropy profile model (see Section 5.2.1). The posterior parameter distributions obtained from the different anisotropy models are then combined by means of the Bayesian model averaging technique, effectively marginalizing over the uncertainties in the orbital anisotropy (see discussion in Section 5.2.2).

5.4.2 Systematic Effects

This section is dedicated to estimating the systematic errors entering our analysis and their impact on the best fit parameter uncertainties.

Malmquist bias

The sample analyzed in this work, being flux-limited, is affected by the Malmquist bias (Sandage, 2000). To estimate the impact of this systematic bias on our results, we estimate its effects on a mock sample. Starting from the halo mass function (Tinker et al., 2008), we create a large mock catalogue, computing the number of expected clusters as a function of halo mass and redshift ($\sim 3.6 \times 10^6$ clusters). We then draw a Poisson realization of this dataset, obtaining a mass selected cluster sample ($4,444$ clusters and $\sim 177,700$ member galaxies) with $M_{200c} \geq 5 \times 10^{13}$ and $0.05 \leq z \leq 0.66$. For each of these clusters we calculate the X-ray luminosity predicted by our scaling relation obtained without the Malmquist bias correction. We then assign to the cluster a luminosity sampled from a Gaussian distribution having the predicted X-ray luminosity as mean, and scatter given by equation (5.4). Here, ΔL_X is drawn from the observed distribution of observed L_X uncertainties as a function of L_X and redshift, extrapolated at lower luminosity. The intrinsic scatter is fixed to the value recovered before the bias correction. The produced mock sample is characterized by $L_X > 1.2 \times 10^{42}$. To create the sample of member galaxies for each cluster, we run MAMPOSSt on a grid of velocities and radii, fixing the galaxy

Number of cluster member galaxies	A_X	B_X	γ_X
$N_{mem} \geq 1$	$0.59^{+0.04}_{-0.05}$	$2.33^{+0.19}_{-0.16}$	$-1.04^{+0.78}_{-0.85}$
$N_{mem} \geq 3$	$0.59^{+0.04}_{-0.04}$	$2.35^{+0.19}_{-0.15}$	$-1.18^{+0.73}_{-0.82}$
$N_{mem} \geq 5$	$0.58^{+0.04}_{-0.05}$	$2.39^{+0.19}_{-0.17}$	$-1.60^{+0.87}_{-0.89}$
$N_{mem} \geq 10$	$0.57^{+0.04}_{-0.05}$	$2.47^{+0.23}_{-0.19}$	$-2.57^{+1.06}_{-1.17}$
$N_{mem} \geq 15$	$0.54^{+0.05}_{-0.05}$	$2.50^{+0.24}_{-0.20}$	$-2.38^{+1.18}_{-1.30}$

Table 5.3: Impact of the number of member galaxies on the luminosity-mass-redshift scaling relation parameters, defined in equation (5.3). The uncertainties on the results are statistical, corresponding to 68 per cent confidence intervals.

number density profile to that described in Section 5.3.6. From the likelihood we derive the probability density of observing an object at a certain projected phase space location (see equation 11, Mamon et al., 2013), drawing a random number of galaxies from the observed distribution of member galaxies.

We fit this mock sample following the procedure described in Section 5.4.1, recovering best fit parameter values consistent with the input values. To finally estimate the effect of the bias, we impose a luminosity cut at $L_X > 4.58 \times 10^{42}$, and fit again. The resulting mass and redshift trends parameters are affected, showing an offset of $\delta B_X = -0.06$ and $\delta\gamma_X = -0.06$. We apply these estimated corrections for the Malmquist bias to all the results we present in this work.

Systematics in MAMPOSSt mass estimates

Another systematic effect we need to take into account is the one associated with the dynamical mass measurements themselves. To estimate this additional systematic uncertainty we employ the findings of Mamon et al. (2013), recovered by analyzing runs of the MAMPOSSt code on numerical simulations. Using particles lying within a sphere of r_{100} around the halo center, they show that the estimated value of the cluster virial radius r_{200c} is biased at $\leq 3.3\%$ (see Table 2, Mamon et al., 2013). We thus adopt a Gaussian systematic uncertainty on the virial mass M_{200c} of $\sigma = 10\%$. As the Mamon et al. (2013) analysis does not explore mass or redshift trends in these biases, we apply the entire uncertainty to the normalization parameter A_X. In a future analysis, we plan to explore the mass and redshift dependence of the systematic uncertainties in dynamical mass estimates from a Jeans analysis on numerical simulations.

Impact of the number of member galaxies

As mentioned in Section 5.3.4, we apply a cut on the number of spectroscopic members per cluster, $N_{mem} \geq 10$. This choice derives from the concern that below a certain number of cluster members even the mean redshift of the cluster becomes uncertain and the dynamical

information becomes too noisy to be reliable for a scaling relation reconstruction. Following C18, we estimate the impact of this cut on our results.

Table 5.3 lists the constraints on the best fit parameters for varying values of $N_{\rm mem}$, from 1 to 15. We note that, as the BCG has already been excluded, clusters with $N_{\rm mem} = 1$ actually have two measured spectroscopic redshifts. As in C18, this cut does not significantly affect the normalization A_X and the mass trend parameter B_X. On the other hand, the redshift trend parameter γ_X changes considerably. The value of γ_X starts converging when including only clusters with at least 10 spectroscopic members, justifying the cut imposed on our sample.

The strong dependence of γ_X on the number of galaxies could indicate an additional source of systematic uncertainty. In the discussion presented in C18, we highlight that the distribution of clusters with $N_{\rm mem} < 10$ extends to higher redshifts, representing a qualitatively different population of objects. To assess whether the trend in γ_X represents the true redshift trend or is a sign of a systematic in the limit of low spectroscopic sampling, further exploration with a larger high z spectroscopic sample is needed. In the meantime, we use this apparent trend to estimate a systematic uncertainty on the scaling relation parameters. As in C18, we define this uncertainty as half the full range of variation in the value of the parameter, $\sigma_{\rm sys,\gamma_X} = \frac{\Delta|\gamma_X|}{2} = 0.76$. We also estimate this factor for the mass trend parameter, $\sigma_{\rm sys,B_X} = \frac{\Delta|B_X|}{2} = 0.08$. For the amplitude parameter, on the other hand, the shift is small compared to the 10% systematic uncertainty described at the beginning of this section. These systematic uncertainties are included in results listed in Table 5.2.

5.4.3 Parameter constraints

The resulting posteriors of our scaling relation parameters are summarized in Table 5.2. The uncertainties are statistical, together with the additional 10% systematic uncertainty described above. Figure 5.3 shows the corresponding joint parameter constraints. We find that galaxy clusters with mass $M_{200c} = 3 \times 10^{14} M_\odot$ at $z = 0.16$ have a mean X-ray luminosity $L_X = 0.57^{+0.04}_{-0.05} \times 10^{44}$, and scale with mass and redshift as $B_X = 2.47^{+0.23}_{-0.19}$ and $\gamma_X = -2.57^{+1.06}_{-1.17}$ respectively. The posterior distribution of the intrinsic scatter is consistent with that of the prior. In the following section we compare our calibration of the L_X-mass relation to previous results from the literature.

5.4.4 Comparison to previous results

Table 5.2 summarizes the parameter estimates and uncertainties for our analysis and the Bulbul et al. (2019) comparison results. To make this comparison, we scale the measurements from Bulbul et al. (2019) to the redshift $z_{\rm piv} = 0.16$, and mass $M_{\rm piv} = 3 \times 10^{14} M_\odot$, using the published best fit redshift and mass trends. The mass conversion from M_{500c} to M_{200c} is carried out using Colossus, an open-source python package for calculations related to cosmology (Diemer, 2017). Moreover, as the analysis performed in Bulbul et al. (2019) is based rest-frame [0.5-2] keV luminosities, we estimate a factor 1.6 to be applied to their amplitude.

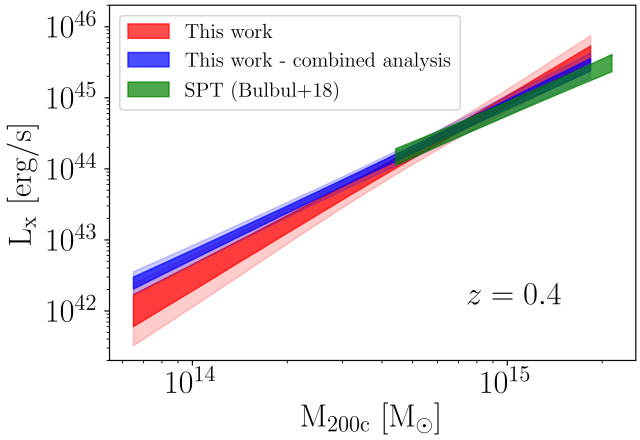

Figure 5.4: Best fit model for our X-ray luminosity-mass relation (in red), evaluated at the redshift $z = 0.4$, compared to the Bulbul et al. (2019) measurements. In blue we show the results from a combined analysis of the two results. Shaded regions correspond to the 1σ and 2σ confidence regions. For the Bulbul et al. (2019) results we only show the 1σ confidence area.

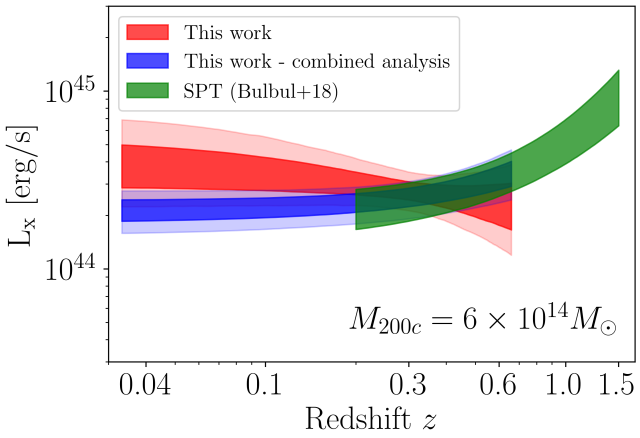

Figure 5.5: Best fit model for our X-ray luminosity-mass relation (in red), evaluated at the mass $M_{200c} = 6 \times 10^{14} M_\odot$, compared to the Bulbul et al. (2019) measurements. In blue we show the results from a combined analysis of the two results. Shaded regions correspond to the 1σ and 2σ confidence regions. For the Bulbul et al. (2019) results we only show the 1σ confidence area.

Figures 5.4 and 5.5 show the mass and redshift trends of the X-ray luminosity, where for the redshift trend we correct the data points to the mass $M_{200c} = 6 \times 10^{14} M_\odot$ and for the mass trend we move the data points to the redshift $z = 0.4$. These values have been chosen as clusters with such mass at this redshift are present in both our dataset and the one analyzed by Bulbul et al. (2019). This means that these are not the places where our constraints are tightest, but the ones where the comparison between the works is justifiable. The best fit model for the $L_X - M_{200c}$ relation is shown in red, with shaded 1σ and 2σ confidence regions. For the results from Bulbul et al. (2019, in green), we show

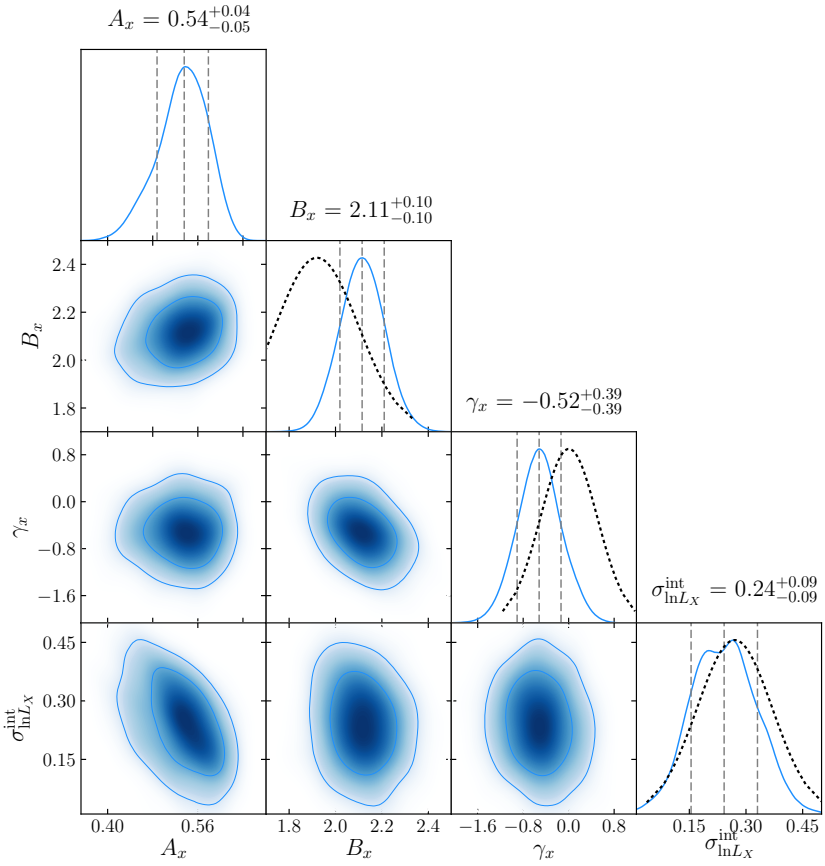

Figure 5.6: Posterior distribution of the scaling relation parameters for the combined analysis. Contours show the 1σ, 2σ, and 3σ confidence regions. The prior distributions for the mass and redshift trends and the intrinsic scatter are shown in black. The prior on the normalization is the flat one showed in Table 5.1.

only the 1σ confidence region. We limit the redshift range to that analyzed in each work.

Our mass trend shows good agreement with the results obtained by Bulbul et al. (2019), based on XMM-*Newton* X-ray observations of an SZE selected sample from the South Pole Telescope 2500 deg^2 SPT-SZ survey (Fig. 5.4). We also find good agreement with the results reported from weak-lensing derived masses of an X-ray selected sample (APEX-SZ; Nagarajan et al., 2018). Additionally, our mass trend is consistent with that found by Giles et al. (2017), obtained through an analysis of galaxy clusters observed with *Chandra*. However, we find a steeper B_X compared to that reported in Mantz et al. (2016) and Vikhlinin et al. (2009), both based on *Chandra* observations of samples derived from the ROSAT All-Sky survey. The former results include weak lensing based masses, while the latter use hydrostatic masses. Overall, our study recovers a steeper than self-similar mass trend, in agreement with most previously published analyses.

Our constraint on the redshift trend of the $L_X - M_{200c} - z$ relation, on the other hand, suggests stronger negative evolution than that found by Bulbul et al. (2019) (Fig. 5.5).

Our redshift trend is also in good agreement with the value of γ_X found by Mantz et al. (2016). We also note that all the results from literature we cited assume a self-similar evolution of the form $E(z)^2$, apart from Mantz et al. (2016) and Vikhlinin et al. (2009). For a discussion of the expected self-similar trends in mass and redshift, we refer the reader to Bulbul et al. (2019).

5.4.5 Combined analysis

Our sample and the one analyzed by Bulbul et al. (2019) cover complementary ranges of mass and redshift. In particular, the SPT selected cluster sample extends to higher redshift, and is therefore helpful in constraining the redshift evolution parameter of the scaling relation. Therefore, we perform a "combined" analysis by adopting the priors on the mass and redshift trends found by Bulbul et al. (2019). Figure 5.6 shows the posterior distribution of the scaling relation parameters, together with the prior distributions for the mass and redshift trends and the intrinsic scatter. The prior on the normalization is the same flat prior used for the CODEX-only analysis (see Table 5.1). Results are also listed in Table 5.2. Figures 5.4, 5.5 and 5.6 all demonstrate that the results of the combined analysis are fully consistent with the Bulbul et al. (2019) ones.

5.5 Conclusions

We present the calibration of the X-ray luminosity–mass–redshift relation using galaxy dynamical information from a sample of 344 CODEX galaxy clusters. These systems are X-ray selected clusters from RASS that have red-sequence selected redMaPPer optical counterparts within a search radius of 3′. The sample is cleaned of random superpositions using an $f_{\rm cont} = 0.05$ cut (Klein et al., 2019), which reduces the contamination from an initial ∼25% to a target 5%. The cluster sample we analyze has redshifts up to $z \sim 0.66$, optical richness $\lambda \geq 20$, and spans an X-ray luminosity range $4.5\times10^{42} \leq L_X/(\mathrm{erg\,s^{-1}}) \leq 3.2\times10^{45}$. The spectroscopic follow-up has been obtained from the SPectroscopic IDentification of eRosita Sources (SPIDERS) survey, resulting in a final sample of 6,658 red member galaxies.

We perform a a Jeans analysis based on the code MAMPOSSt (Mamon et al., 2013). For each individual cluster, we extract the likelihood of consistency between the projected phase space distribution of the cluster members with measured redshifts and the modeled projected distribution for a cluster at redshift z, luminosity L_X, and inferred mass M_{200c}. We adopt an NFW profile for the red galaxy tracer population with concentration $c = 5.37$ (Hennig et al., 2017, and Section 5.3.6), and employ five different velocity dispersion anisotropy profiles. We combine luminosity-mass relation posterior parameter distributions from the different anisotropy models by performing Bayesian model averaging, allowing us to marginalize over the orbital anisotropy of the spectroscopic galaxy population.

The scaling relation is modeled as $L_X \propto A_X M_{200c}^{B_X} E(z)^2 (1+z)^{\gamma_X}$ (equation 5.3). We correct for the Eddington bias by implementing the method described in Mortonson et al.

(2011), which provides an estimate of the mean mass shift due to the log-normal mass observable relation scatter (equation 5.4) together with the measurement uncertainties on the X-ray luminosity. We also correct for the Malmquist bias, after evaluating its effect on a mock sample.

Results are showed in Table 5.2. For clusters of mass $M_{\rm piv} = 3 \times 10^{14} M_{\odot}$, at redshift $z_{\rm piv} = 0.16$, we find the following constraints on the scaling relation parameters:

$$\begin{aligned} \mathrm{A_X} =& 0.57^{+0.04}_{-0.05} \pm 0.06, \\ \mathrm{B_X} =& 2.47^{+0.23}_{-0.19} \pm 0.08, \\ \gamma_{\rm X} =& -2.57^{+1.06}_{-1.17} \pm 0.76. \end{aligned} \tag{5.7}$$

where we quote systematic uncertainties for all the parameters. The amplitude uncertainty of 10% comes from an estimate of the dynamical mass systematic uncertainty, applied to the scaling relation amplitude $\mathrm{A_X}$ (see study of systematics in Mamon et al., 2013).

Our results on the mass trend of the scaling relation are steeper, but statistically consistent (within 2σ) with some previous literature results (Bulbul et al., 2019; Nagarajan et al., 2018; Giles et al., 2017). However, we find large departures from the Mantz et al. (2016) and Vikhlinin et al. (2009) mass trends.

We examine the redshift trend of the $L_{\rm X}$–mass scaling relation, finding a stronger negative, non-self-similar evolution of $L_{\rm X}$ with redshift with respect to the Bulbul et al. (2019) results. We explore this result by performing our analysis fixing the priors on our mass and redshift slopes to those adopted in Bulbul et al. (2019). We recover the following set of parameters:

$$\begin{aligned} \mathrm{A_X} =& 0.54^{+0.04}_{-0.05}, \\ \mathrm{B_X} =& 2.11^{+0.10}_{-0.10}, \\ \gamma_{\rm X} =& -0.52^{+0.39}_{-0.39}. \end{aligned} \tag{5.8}$$

We note that the redshift trend has shifted to significant higher values, being consistent with the self-similar evolution and with previous studies.

This work, together with C18, shows the potential of dynamical masses in deriving mass–observable relations even in the limit of small number of cluster members. This very promising result will be extremely useful in the context of future spectroscopic surveys like DESI (Levi et al., 2013), 4MOST (de Jong et al., 2012) and Euclid (Laureijs et al., 2011). Performing a dynamical analysis on numerical simulations will enable significant improvements in the assessment of further systematic uncertainties, such as the impact of residual interlopers in our sample, departures from virial equilibrium, and variation of the velocity anisotropy profile (Capasso et al., in prep.).

Acknowledgments

The Munich group acknowledges the support by the DFG Cluster of Excellence "Origin and Structure of the Universe", the MPG faculty fellowship program and the Ludwig-Maximilians-Universität Munich. RC acknowledges participation in the IMPRS on Astrophysics at the Ludwig-Maximilians University and the associated financial support from the Max-Planck Society. RC and VS acknowledge support from the German Space Agency (DLR) through *Verbundforschung* project ID 50OR1603. AS is supported by the ERC-StG "ClustersXCosmo", grant agreement 716762. AB acknowledges the hospitality of the LMU and partial financial support from PRIN-INAF 2014 "Glittering kaleidoscopes in the sky: the multifaceted nature and role of Galaxy Clusters?", P.I.: Mario Nonino.
Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.
SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Chapter 6

Summary and outlook

The analyses presented in this thesis are motivated by the need to calibrate cluster masses with high precision. We exploit the information residing in the projected phase-space (line of sight velocity, projected distance from cluster center) of cluster galaxies to recover the cluster gravitational potential, while gaining insight into the formation history of the cluster and the evolution of its member galaxies. In this final chapter, we summarize the results of these analyses and give an outlook for future developments and related work we intend to pursue in the future.

In Chapter 3 we analyze a sample of ∼3000 passive, non-emission line cluster galaxies drawn from 110 SZE selected galaxy clusters identified within the SPT-SZ 2500 deg^2 survey. We exploit the full information residing in the projected phase-space distributions of these clusters by performing a dynamical Jeans analysis.

We study the orbital characteristics of the passive galaxies, finding velocity dispersion anisotropy profiles that are approximately isotropic near the center and increasingly radial toward the virial radius. Interestingly, this distribution does not depend on cluster mass or redshift, providing an indication that the accretion of material from the surround large scale structure is occurring similarly at low and high redshift.

We then study the pseudo-phase-space density profile. The PPSD profiles show that the analyzed sample is constituted by approximately virialized halos that have undergone violent relaxation, because the orbital properties of the passive galaxy sample is in good agreement with that of N-body particles within clusters forming in structure formation simulations. Moreover, we find that the amplitude of the profile scales with redshift and mass as expected under the assumption of self-similarity. Given the wide redshift range of our dataset, $0.2 < z < 1.3$, we explore eventual trends with mass and redshift for the anisotropy and PPSD profiles. The absence of a redshift trend in both profiles suggests that the process of infall and relaxation for the passive galaxy population is occurring similarly at all redshifts since at least $z \sim 1$.

We compare dynamical masses with those obtained through independent calibration in previous analyses. The dynamical mass constraints are in good agreement with external

mass estimates of the SPT cluster sample from either weak lensing, velocity dispersions, or X-ray Y_X measurements. However, the dynamical analysis suggests lower masses than required by the preferred Planck CMB anisotropy ΛCDM model. The discrepancy grows with redshift, where in the highest redshift bin the ratio of dynamical to SPT+Planck masses is $\eta = 0.63^{+0.13}_{-0.08} \pm 0.06$ (statistical and systematic), corresponding to a 2.6σ discrepancy.

Finally, using *Chandra* X-ray data, we examine the impact of the dynamical state of the clusters on our dynamical analysis. We find dynamical masses to be in good agreement with our combined sample for both the relaxed and un-relaxed clusters. However, we find large uncertainties on the anisotropy profiles, especially for the relaxed sample, allowing also for isotropic orbits at all radii. Further investigation with larger dynamical samples in cluster ensembles is thus necessary to understand the different behaviour of these objects.

As a next step, this analysis can be extended to datasets containing lower-mass systems, such as the sample being built up through spectroscopic observations of optically selected clusters within the Dark Energy Survey. On the longer term, we expect the EUCLID survey, together with spectroscopic followup of the LSST survey, to provide larger galaxy samples that include both star forming and passive galaxies. These samples will allow cluster masses to be constrained in a redshift regime where weak lensing is challenging, while also enabling studies of the kinematic relationship between cluster star forming and passive galaxies out to higher redshifts.

In Chapter 4 we adapt the previous dynamical analysis to perform a richness-mass-redshift scaling relation calibration. We employ a sample of 428 X-ray selected CODEX galaxy clusters. These are RASS systems that have red sequence selected redMaPPer optical counterparts within a search radius of $3'$. Our sample extends to $z \sim 0.66$ and is characterized by optical richness $\lambda \geq 20$. The spectroscopic follow-up obtained from the SPIDERS survey provides us with 7807 red member galaxies for our analysis.

We model the scaling relation as $\lambda \propto \mathrm{A}_\lambda {M_{200c}}^{\mathrm{B}_\lambda}(1+z)^{\gamma_\lambda}$. After applying corrections for the Eddington bias and for the Malmquist bias, we find constraints on the scaling relation parameters to be as follows: the normalization$\mathrm{A}_\lambda = 38.6^{+3.1}_{-4.1} \pm 3.9$, mass slope $\mathrm{B}_\lambda = 0.99^{+0.06}_{-0.07} \pm 0.04$ and redshift slope $\gamma_\lambda = -1.13^{+0.32}_{-0.34} \pm 0.49$. We present systematic uncertainties as a second component.

Our results on the mass trends are in good agreement with previously published analyses, with the exception of those from stacked weak lensing. This discrepancy could be due to differences in the levels of contamination between pure redMaPPer and RASS+redMaPPer samples, such as our CODEX sample. The redshift trend we measure is more negative than, but statistically consistent with, previous results. Our trend reflects a change in the cluster galaxy red sequence fraction with redshift that is consistent with direct measurements. We also examine the impact of a correlated scatter in X-ray luminosity and optical richness, showing it has negligible impact on our results.

In Chapter 5 we extend the work on the mass–observable scaling relation calibration to X-ray luminosities. In contrast to our study of the richness–mass scaling relation, for this

study we have to take into account the fact that X-ray sources are likely to be contaminated by random superpositions of non-cluster X-ray candidates with optical systems. We assign each source a probability of being a random superposition, and produce a sample with a 5% contamination fraction, which is a factor of ∼5 improvement over the base CODEX sample. We obtain a sample of 344 CODEX clusters up to $z \sim 0.66$, consisting of ∼6600 red member galaxies, optical richnesses $\lambda \geq 20$, and spanning an X-ray luminosity range $4.5 \times 10^{42} \leq L_{\rm X}/({\rm erg\,s^{-1}}) \leq 3.2 \times 10^{45}$. We model the scaling relation as $L_{\rm X} \propto {\rm A_X} M_{\rm 200c}^{\rm B_X} E(z)^2 (1+z)^{\gamma_{\rm X}}$. As before, we correct for the Eddington and Malmquist biases. We find the parameter constraints to be ${\rm A_X} = 0.57^{+0.04}_{-0.05} \times 10^{44}$ erg s^{-1}, ${\rm B_X} = 2.47^{+0.23}_{-0.19}$, $\gamma_{\rm X} = -2.57^{+1.06}_{-1.17}$.

Our results on the mass trend of the scaling relation are steeper, but statistically consistent with most previous literature results. Our redshift trend, on the other hand, presents a very uncertain constraint that suggests stronger negative evolution of $L_{\rm X}$ with redshift with respect to the Bulbul et al. (2018) results. To improve the constraint on the redshift trend of the scaling relation, we perform a combined analysis, fixing the priors on our mass and redshift slopes to the posterior parameter distributions reported in Bulbul et al. (2018). We recover the following set of parameters: ${\rm A_X} = 0.54^{+0.04}_{-0.05} \times 10^{43}$ erg s^{-1}, ${\rm B_X} = 2.11^{+0.10}_{-0.10}$, $\gamma_{\rm X} = -0.52^{+0.39}_{-0.39}$. We note that the redshift trend has shifted to significant higher values, and is consistent with the self-similar evolution and with previous studies. With a larger dynamical sample extending to higher redshift in the future, we could test this redshift evolution more directly.

The analyses and calibrations of scaling relations show that dynamical masses are a powerful tool to gain information on the link between the masses of galaxy clusters and readily obtainable observables– even in the limit of large cluster samples with small spectroscopic samples available for each cluster.

Performing a dynamical analysis on numerical simulations will enable significant improvements in the assessment of further systematic uncertainties, such as the impact of residual interlopers in our sample, departures from equilibrium, and variation of the velocity anisotropy profile (Capasso et al., in prep.). With this purpose in mind, we developed a fully parallelized Python code, including Bayesian statistical inference methods and model averaging techniques. This code allows a greater flexibility on the assumed anisotropy and mass profiles, and an easier coupling of dynamical mass-calibration of galaxy clusters within a framework for cluster number-counts cosmology experiments.

Our analysis demonstrates that there is promise in the analysis of small per-cluster spectroscopic samples of the sort that will be delivered by future spectroscopic surveys like DESI (Levi et al., 2013), 4MOST (de Jong et al., 2012) and Euclid (Laureijs et al., 2011).

Finally, we plan on combining high-precision kinematic and lensing measurements of the total mass profile of galaxy clusters to estimate the value of anisotropic stress, i.e. the ratio $\eta = \Phi/\Psi$ between the two scalar potentials in the linear perturbed Friedmann-Lemaitre-Robertson-Walker metric. Dynamical and weak lensing mass profiles are connected to the gravitational potentials in different ways: the first probes the motion of galaxies in clusters, sensitive only to Φ, while the second, studying the paths of photons, perceives the contribution of both potentials. An accurate measurement of this ratio could show

possible, interesting deviations from the predictions of the theory of General Relativity, according to which, in the absence of relativistic sources, Φ is equal to Ψ.

Bibliography

Abell G. O., 1958, ApJS, 3, 211

Adami C., Biviano A., Mazure A., 1998, A&A, 331, 439

Aguerri J. A. L., Agulli I., Diaferio A., Dalla Vecchia C., 2017, MNRAS, 468, 364

Aguilar M. et al., 2013, Physical Review Letters, 110, 141102

Aihara H. et al., 2011, ApJS, 193, 29

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARAA, 49, 409

Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879

Allen S. W., Schmidt R. W., Fabian A. C., 2002, MNRAS, 334, L11

Allen S. W., Schmidt R. W., Fabian A. C., Ebeling H., 2003, MNRAS, 342, 287

Allington-Smith J. et al., 1994, PASP, 106, 983

Amodeo S., Ettori S., Capasso R., Sereno M., 2016, A&A, 590, A126

Andersson K. et al., 2011, ApJ, 738, 48

Andreon S., 1996, A&A, 314, 763

Angloher G. et al., 2012, European Physical Journal C, 72, 1971

Annis J. et al., 2014, ApJ, 794, 120

Annunziatella M. et al., 2016, A&A, 585, A160

Appenzeller I. et al., 1998, The Messenger, 94, 1

Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&A, 474, L37

Aslanyan G., Price L. C., Abazajian K. N., Easther R., 2014, J. Chem. Phys., 8, 052

Austin C. G., Williams L. L. R., Barnes E. I., Babul A., Dalcanton J. J., 2005, ApJ, 634, 756

Bahcall J. N., Tremaine S., 1981, ApJ, 244, 805

Bahcall N. A., Fan X., 1998, ApJ, 504, 1

Bahcall N. A. et al., 2003, ApJS, 148, 243

Barnes D. J. et al., 2017, MNRAS, 471, 1088

Barnes E. I., Williams L. L. R., Babul A., Dalcanton J. J., 2006, ApJ, 643, 797

Bartelmann M., 2010, Classical and Quantum Gravity, 27, 233001

Baxter E. J. et al., 2018, MNRAS, 476, 2674

Baxter E. J., Rozo E., Jain B., Rykoff E., Wechsler R. H., 2016, MNRAS, 463, 205

Bayliss M. B. et al., 2016, ApJS, 227, 3

Bayliss M. B. et al., 2017, ApJ, 837, 88

Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Bellovary J. M., Dalcanton J. J., Babul A., Quinn T. R., Maas R. W., Austin C. G., Williams L. L. R., Barnes E. I., 2008, ApJ, 685, 739

Benson B. A. et al., 2013, ApJ, 763, 147

Bertschinger E., 1985, ApJS, 58, 39

Biffi V. et al., 2016, ApJ, 827, 112

Binney J., Mamon G. A., 1982, MNRAS, 200, 361

Binney J., Tremaine S., 1987, Galactic dynamics. Princeton, NJ, Princeton University Press, 1987, 747 p.

Biviano A., 2000, in Constructing the Universe with Clusters of Galaxies

Biviano A., Girardi M., 2003, ApJ, 585, 205

Biviano A., Girardi M., Giuricin G., Mardirossian F., Mezzetti M., 1992, ApJ, 396, 35

Biviano A., Katgert P., 2004, A&A, 424, 779

Biviano A., Katgert P., Mazure A., Moles M., den Hartog R., Perea J., Focardi P., 1997, A&A, 321, 84

Biviano A., Katgert P., Thomas T., Adami C., 2002, A&A, 387, 8

Biviano A., Murante G., Borgani S., Diaferio A., Dolag K., Girardi M., 2006, A&A, 456, 23

Biviano A., Poggianti B. M., 2009, A&A, 501, 419

Biviano A., Popesso P., Dietrich J. P., Zhang Y.-Y., Erfanianfar G., Romaniello M., Sartoris B., 2017, A&A, 602, A20

Biviano A. et al., 2013, A&A, 558, A1

Biviano A., van der Burg R. F. J., Muzzin A., Sartoris B., Wilson G., Yee H. K. C., 2016, A&A, 594, A51

Blanton M. R. et al., 2017, AJ, 154, 28

Bleem L. E. et al., 2015, ApJS, 216, 27

Bocquet S. et al., 2018, arXiv e-prints

Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361

Bocquet S. et al., 2015, ApJ, 799, 214

Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603

Bondi H., Gold T., 1948, MNRAS, 108, 252

Borgani S. et al., 2001, ApJ, 561, 13

Bower R. G., Terlevich A., Kodama T., Caldwell N., 1999, in Astronomical Society of the Pacific Conference Series, Vol. 163, Star Formation in Early Type Galaxies, Carral P., Cepa J., eds., p. 211

Briel U. G., Henry J. P., Boehringer H., 1992, A&A, 259, L31

Bulbul E. et al., 2018, ArXiv e-prints

Bulbul E. et al., 2019, ApJ, 871, 50

Buote D. A., Tsai J. C., 1995, ApJ, 452, 522

Buote D. A., Tsai J. C., 1996, ApJ, 458, 27

Burkert A., 1995, ApJ, 447, L25

Capasso R. et al., 2018, arXiv e-prints

Capasso R. et al., 2019, MNRAS, 482, 1043

Carlberg R. G., Yee H. K. C., Ellingson E., 1997a, ApJ, 478, 462

Carlberg R. G. et al., 1997b, ApJ, 485, L13

Carlberg R. G. et al., 1997c, ApJ, 476, L7

Carlstrom J. E. et al., 2011, PASP, 123, 568

Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARAA, 40, 643

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Ceccarelli M. L., Valotto C., Lambas D. G., Padilla N., Giovanelli R., Haynes M., 2005, ApJ, 622, 853

Chandrasekhar S., 1943, ApJ, 97, 255

Chincarini G., Rood H. J., 1977, ApJ, 214, 351

Chiu I. et al., 2016a, MNRAS, 457, 3050

Chiu I. et al., 2016b, MNRAS, 455, 258

Clerc N. et al., 2016, MNRAS, 463, 4490

Clerc N. et al., 2018, A&A, 617, A92

Coles P., Lucchin F., 2002, Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition. p. 512

Corless V. L., King L. J., 2009, MNRAS, 396, 315

Costanzi M. et al., 2018, ArXiv e-prints

Cypriano E. S., Sodré, Jr. L., Kneib J.-P., Campusano L. E., 2004, ApJ, 613, 95

Dalton G. B., Maddox S. J., Sutherland W. J., Efstathiou G., 1997, MNRAS, 289, 263

Dawson K. S. et al., 2016, AJ, 151, 44

de Haan T. et al., 2016, ApJ, 832, 95

de Jong R. S. et al., 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series., Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, p. 84460T

de Propris R., Stanford S. A., Eisenhardt P. R., Dickinson M., Elston R., 1999, AJ, 118, 719

de Sitter W., 1917, MNRAS, 78, 3

Dehnen W., McLaughlin D. E., 2005, MNRAS, 363, 1057

den Hartog R., Katgert P., 1996, MNRAS, 279, 349

Diaferio A., 1999, MNRAS, 309, 610

Diaferio A., Geller M. J., 1997, ApJ, 481, 633

Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859

Diemer B., 2017, ArXiv e-prints

Dietrich J. P. et al., 2019, MNRAS, 483, 2871

Dietrich J. P. et al., 2017, ArXiv: 1711.05344

Donahue M., Voit G. M., Gioia I., Luppino G., Hughes J. P., Stocke J. T., 1998, ApJ, 502, 550

Donahue M. et al., 2014, ApJ, 794, 136

Dressler A., 1980, ApJ, 236, 351

Dressler A., 1984, ARAA, 22, 185

Dressler A., Hare T., Bigelow B. C., Osip D. J., 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, p. 62690F

Dressler A., Shectman S. A., 1988, AJ, 95, 985

Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler, Jr. A., 1999, ApJS, 122, 51

Du N. et al., 2018, Physical Review Letters, 120, 151301

Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64

Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359

Dye S., Taylor A. N., Thommes E. M., Meisenheimer K., Wolf C., Peacock J. A., 2001, MNRAS, 321, 685

Eddington A. S., 1913, MNRAS, 73, 359

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Einstein A., 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.

Einstein A., 1916, Annalen der Physik, 354, 769

Einstein A., 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.

Eisenhardt P. R. M. et al., 2008, ApJ, 684, 905

Eisenstein D. J. et al., 2005, ApJ, 633, 560

Ellingson E., Lin H., Yee H. K. C., Carlberg R. G., 2001, ApJ, 547, 609

Ettori S., Donnarumma A., Pointecouteau E., Reiprich T. H., Giodini S., Lovisari L., Schmidt R. W., 2013, SSRv, 177, 119

Evrard A. E. et al., 2008, ApJ, 672, 122

Fabian A. C., 1991, MNRAS, 253, 29P

Fabian A. C., Sanders J. S., Allen S. W., Crawford C. S., Iwasawa K., Johnstone R. M., Schmidt R. W., Taylor G. B., 2003, MNRAS, 344, L43

Fadda D., Girardi M., Giuricin G., Mardirossian F., Mezzetti M., 1996, ApJ, 473, 670

Faltenbacher A., Hoffman Y., Gottlöber S., Yepes G., 2007, MNRAS, 376, 1327

Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, MNRAS, 460, 3900

Farahi A. et al., 2018, A&A, 620, A8

Feroz F., Hobson M. P., 2008, MNRAS, 384, 449

Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601

Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, ArXiv e-prints

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fowler J. W. et al., 2007, Applied Optics, 46, 3444

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Gao L., De Lucia G., White S. D. M., Jenkins A., 2004, MNRAS, 352, L1

Gavazzi R., 2005, A&A, 443, 793

Geller M. J., Beers T. C., 1982, PASP, 94, 421

Geller M. J., Diaferio A., Kurtz M. J., 1999, The Astrophysical Journal Letters, 517, L23

Gerhard O. E., 1993, MNRAS, 265, 213

Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, MNRAS, 300, 146

Giles P. A. et al., 2017, MNRAS, 465, 858

Girardi M., Fadda D., Escalera E., Giuricin G., Mardirossian F., Mezzetti M., 1997, ApJ, 490, 56

Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148

Gladders M. D., Yee H. K. C., Majumdar S., Barrientos L. F., Hoekstra H., Hall P. B., Infante L., 2007, ApJ, 655, 128

Gunn J. E. et al., 2006, AJ, 131, 2332

Guth A. H., 1981, Phys. Rev. D, 23, 347

Haiman Z., Mohr J. J., Holder G. P., 2001, ApJ, 553, 545

Haines C. P. et al., 2018, MNRAS, 477, 4931

Hasselfield M. et al., 2013, JCAp, 7, 008

Hee S., Handley W. J., Hobson M. P., Lasenby A. N., 2016, MNRAS, 455, 2461

Hennig C. et al., 2017, MNRAS, 467, 4015

Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307

Hernquist L., 1990, ApJ, 356, 359

Herschel F. W., 1785, Phil. Trans., 75, 213

Hoekstra H., 2007, MNRAS, 379, 317

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685

Hoessel J. G., Gunn J. E., Thuan T. X., 1980, ApJ, 241, 486

Hoeting J. A., Madigan D., Raftery A. E., Volinsky C. T., 1999, STATISTICAL SCIENCE, 14, 382

Holder G., Haiman Z., Mohr J. J., 2001, ApJ, 560, L111

Holmberg E., 1940, ApJ, 92, 200

Hook I. M., Jørgensen I., Allington-Smith J. R., Davies R. L., Metcalfe N., Murowinski R. G., Crampton D., 2004, PASP, 116, 425

Hoyle F., 1948, MNRAS, 108, 372

Hoyle F., Narlikar J. V., 1963, Proceedings of the Royal Society of London Series A, 273, 1

Hubble E., 1929, Contributions from the Mount Wilson Observatory, vol. 3, pp.23-28, 3, 23

Hubble E., Humason M. L., 1931, ApJ, 74, 43

Hurier G., Angulo R. E., 2018, A&A, 610, L4

Iannuzzi F., Dolag K., 2012, MNRAS, 427, 1024

Jeans J. H., 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Jeffreys H., 1998, The Theory of Probability, Oxford Classic Texts in the Physical Sciences. OUP Oxford

Jimeno P., Broadhurst T., Lazkoz R., Angulo R., Diego J.-M., Umetsu K., Chu M.-c., 2017, MNRAS, 466, 2658

Jullo E., Natarajan P., Kneib J.-P., D'Aloisio A., Limousin M., Richard J., Schimd C., 2010, Science, 329, 924

Katgert P., Biviano A., Mazure A., 2004a, ApJ, 600, 657

Katgert P., Biviano A., Mazure A., 2004b, ApJ, 600, 657

Khullar G. et al., 2018, ArXiv e-prints

Klein M., Grandis S., Mohr J., Paulus M., the DES Collaboration, 2019, arXiv e-prints

Klein M. et al., 2018, MNRAS, 474, 3324

Komatsu E. et al., 2011, The Astrophysical Journal Supplement Series, 192, 18

Kragh H., 1996a, Cosmology and controversy. The historical development of two theories of the universe

Kragh H., 1996b, Centaurus, 38, 335

Kronawitter A., Saglia R. P., Gerhard O., Bender R., 2000, 144, 53

Lapi A., Cavaliere A., 2009, The Astrophysical Journal, 692, 174

Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129

Laureijs R. et al., 2011, ArXiv e-prints

Lemaître G., 1931, MNRAS, 91, 483

Lemze D., Broadhurst T., Rephaeli Y., Barkana R., Umetsu K., 2009, ApJ, 701, 1336

Levi M. et al., 2013, ArXiv e-prints

Lima M., Hu W., 2005, Phys. Rev. D, 72, 043006

Lin Y., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745

Lin Y.-T., Mohr J. J., 2004, ApJ, 617, 879

Lin Y.-T., Mohr J. J., Gonzalez A. H., Stanford S. A., 2006, ApJ, 650, L99

Łokas E. L., 2002, MNRAS, 333, 697

Łokas E. L., Mamon G. A., 2003, MNRAS, 343, 401

Łokas E. L., Wojtak R., Gottlöber S., Mamon G. A., Prada F., 2006, MNRAS, 367, 1463

Lovisari L., Reiprich T. H., Schellenberger G., 2015, A&A, 573, A118

Lumsden S. L., Nichol R. C., Collins C. A., Guzzo L., 1992, MNRAS, 258, 1

Lyapustin D., 2011, arXiv e-prints

Lynden-Bell D., 1967, MNRAS, 136, 101

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008a, MNRAS, 391, 1940

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008b, MNRAS, 391, 1940

Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567

Malmquist K. G., 1920, Meddelanden fran Lunds Astronomiska Observatorium Serie II, 22, 1

Mamon G. A., 1992, ApJ, 401, L3

Mamon G. A., Biviano A., Boué G., 2013, MNRAS, 429, 3079

Mamon G. A., Biviano A., Murante G., 2010a, A&A, 520, A30

Mamon G. A., Biviano A., Murante G., 2010b, A&A, 520, A30

Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705

Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, MNRAS, 406, 1773

Mantz A. et al., 2014, in AAS/High Energy Astrophysics Division, Vol. 14, AAS/High Energy Astrophysics Division #14, p. 102.02

Mantz A. B. et al., 2016, MNRAS, 463, 3582

Mantz A. B. et al., 2015, MNRAS, 446, 2205

McClintock T. et al., 2019, MNRAS, 482, 1352

McDonald M. et al., 2017, ApJ, 843, 28

McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937

McNamara B. R., Nulsen P. E. J., 2007, ARAA, 45, 117

Mei S. et al., 2009, ApJ, 690, 42

Melin J.-B., Bartlett J. G., 2015, A&A, 578, A21

Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93

Meneghetti M. et al., 2014, ApJ, 797, 34

Merloni A. et al., 2012, ArXiv e-prints

Merritt D., 1985, MNRAS, 214, 25P

Merritt D., 1987, ApJ, 313, 121

Merritt D., Graham A. W., Moore B., Diemand J., Terzić B., 2006, AJ, 132, 2685

Messier C., 1784, Connaissance des Temp

Metzler C. A., White M., Loken C., 2001, ApJ, 547, 560

Milne E. A., 1935, Relativity, gravitation and world-structure

Miralda-Escude J., Babul A., 1995, ApJ, 449, 18

Mirkazemi M. et al., 2015, ApJ, 799, 60

Mohr J. J., Evrard A. E., Fabricant D. G., Geller M. J., 1995, ApJ, 447, 8+

Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 413, 492

Mohr J. J., Geller M. J., Fabricant D. G., Wegner G., Thorstensen J., Richstone D. O., 1996, ApJ, 470, 724

Moore B., Katz N., Lake G., 1996, ApJ, 457, 455

Mortonson M. J., Hu W., Huterer D., 2011, Phys. Rev. D, 83, 023015

Moss C., Dickens R. J., 1977, MNRAS, 178, 701

Mukhanov V. F., Chibisov G. V., 1981, ZhETF Pisma Redaktsiiu, 33, 549

Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, MNRAS, 430, 2638

Munari E., Biviano A., Mamon G. A., 2014, A&A, 566, A68

Munari E., Biviano A., Mamon G. A., 2015, A&A, 574, C1

Murata R., Nishimichi T., Takada M., Miyatake H., Shirasaki M., More S., Takahashi R., Osato K., 2018, ApJ, 854, 120

Muzzin A. et al., 2014, ApJ, 796, 65

Muzzin A. et al., 2012, ApJ, 746, 188

Muzzin A., Yee H. K. C., Hall P. B., Ellingson E., Lin H., 2007, ApJ, 659, 1106

Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98

Nagarajan A. et al., 2018, MNRAS

Narayan R., Bartelmann M., 1996, arXiv e-prints, astro

Natarajan P., Kneib J.-P., 1997, MNRAS, 287, 833

Navarro J., Frenk C., White S., 1996, ApJ, 462, 563

Navarro J., Frenk C., White S., 1997, ApJ, 490, 493

Navarro J. F. et al., 2010, MNRAS, 402, 21

Noble A. G., Webb T. M. A., Yee H. K. C., Muzzin A., Wilson G., van der Burg R. F. J., Balogh M. L., Shupe D. L., 2016, ApJ, 816, 48

Nurgaliev D. et al., 2017, ApJ, 841, 5

Nurgaliev D., McDonald M., Benson B. A., Miller E. D., Stubbs C. W., Vikhlinin A., 2013, ApJ, 779, 112

Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, ApJ, 721, 875

Old L. et al., 2015, MNRAS, 449, 1897

Osipkov L. P., 1979, Soviet Astronomy Letters, 5, 42

Ota N., Pointecouteau E., Hattori M., Mitsuda K., 2004, ApJ, 601, 120

Pacaud F. et al., 2016, A&A, 592, A2

Pacaud F. et al., 2007, MNRAS, 382, 1289

Papovich C., 2008, ApJ, 676, 206

Peebles P. J. E., 1971, Physical cosmology

Peebles P. J. E., 1993, Principles of Physical Cosmology

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S. et al., 1999, ApJ, 517, 565

Planck Collaboration et al., 2011, A&A, 536, A13

Planck Collaboration et al., 2014a, A&A, 571, A16

Planck Collaboration et al., 2014b, A&A, 571, A20

Planck Collaboration et al., 2016a, A&A, 594, A17

Planck Collaboration et al., 2016b, A&A, 594, A20

Planck Collaboration et al., 2016c, A&A, 594, A13

Planck Collaboration et al., 2016d, A&A, 594, A24

Planck Collaboration et al., 2018a, arXiv e-prints, arXiv:1807.06205

Planck Collaboration et al., 2018b, arXiv e-prints

Postman M., Lubin L. M., Gunn J. E., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, AJ, 111, 615

Powell, 2006, in Large-Scale Nonlinear Optimization, Springer US

Pratt G. W., Arnaud M., 2002, A&A, 394, 375

Pratt G. W., Böhringer H., Croston J. H., Arnaud M., Borgani S., Finoguenov A., Temple R. F., 2007, A&A, 461, 71

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361

Predehl P. et al., 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series., Vol. 7732, Space Telescopes and Instrumentation 2010: Ultraviolet to Gamma Ray, p. 77320U

Press W. H., Schechter P., 1974, ApJ, 187, 425

Raine D., 1975, Monthly Notices of the Royal Astronomical Society, 171, 507

Rasia E. et al., 2006, MNRAS, 369, 2013

Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716

Riess A. G. et al., 1998, AJ, 116, 1009

Rines K., Diaferio A., 2006, AJ, 132, 1275

Rines K., Geller M. J., Kurtz M. J., Diaferio A., 2003, AJ, 126, 2152

Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, ApJ, 819, 63

Rines K. J., Geller M. J., Diaferio A., Hwang H. S., Sohn J., 2018, ApJ, 862, 172

Robertson H. P., 1935, ApJ, 82, 284

Rood H. J., 1965, PhD thesis, THE UNIVERSITY OF MICHIGAN.

Rood H. J., Page T. L., Kintner E. C., King I. R., 1972, ApJ, 175, 627

Roszkowski L., Sessolo E. M., Trojanowski S., 2018, Reports on Progress in Physics, 81, 066201

Ruel J. et al., 2014, ApJ, 792, 45

Rykoff E. S. et al., 2012, ApJ, 746, 178

Rykoff E. S. et al., 2014, ApJ, 785, 104

Rykoff E. S. et al., 2016, ApJS, 224, 1

Sandage A., 2000, Malmquist Bias and Completeness Limits, Murdin P., ed., p. 1940

Sarazin C. L., 1988, X-ray emission from clusters of galaxies

Saro A. et al., 2017, MNRAS, 468, 3347

Saro A. et al., 2015, MNRAS, 454, 2305

Saro A., Mohr J. J., Bazin G., Dolag K., 2013, ApJ, 772, 47

Schaffer K. K. et al., 2011, ApJ, 743, 90

Schechter P., 1976, ApJ, 203, 297

Schrabback T. et al., 2018, MNRAS, 474, 2635

Sereno M., Ettori S., 2017, MNRAS, 468, 3322

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Shi X., Komatsu E., 2014, MNRAS, 442, 521

Sifón C. et al., 2016, MNRAS, 461, 248

Sifón C. et al., 2013a, ApJ, 772, 25

Sifón C. et al., 2013b, ApJ, 772, 25

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, MNRAS, 466, 3103

Smoot G. F. et al., 1992, ApJ, 396, L1

Soergel B. et al., 2016, MNRAS, 461, 3172

Sohn J., Geller M. J., Rines K. J., Hwang H. S., Utsumi Y., Diaferio A., 2018, ApJ, 856, 172

Solanes J. M., Salvador-Sole E., 1990, A&A, 234, 93

Song J., Mohr J. J., Barkhouse W. A., Warren M. S., Dolag K., Rude C., 2012a, ApJ, 747, 58

Song J. et al., 2012b, ApJ, 761, 22

Spergel D. N. et al., 2003, The Astrophysical Journal Supplement Series, 148, 175

Springel V. et al., 2005, Nature, 435, 629

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Staniszewski Z. et al., 2009, ApJ, 701, 32

Stein P., 1997, A&A, 317, 670

Stern C. et al., 2018, ArXiv e-prints

Strazzullo V. et al., 2018, ArXiv e-prints

Strigari L. E., Bullock J. S., Kaplinghat M., Simon J. D., Geha M., Willman B., Walker M. G., 2008, Nature, 454, 1096

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Tammann G. A., 1972, A&A, 21, 355

Tarter J., Silk J., 1974, Quarterly Journal of the Royal Astronomical Society, 15, 122

Taylor A. N., Kitching T. D., Bacon D. J., Heavens A. F., 2007, MNRAS, 374, 1377

Taylor J. E., Navarro J. F., 2001, ApJ, 563, 483

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tiret O., Combes F., Angus G. W., Famaey B., Zhao H. S., 2007, A&A, 476, L1

Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Coe D., Postman M., 2011, ApJ, 738, 41

Umetsu K. et al., 2014, ApJ, 795, 163

Van den Bergh S., 1969, Nature, 224, 891

van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015a, A&A, 577, A19

van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015b, A&A, 577, A19

van der Burg R. F. J., Muzzin A., Hoekstra H., Wilson G., Lidman C., Yee H. K. C., 2014, A&A, 561, A79

van der Marel R. P., 1994, MNRAS, 270, 271

van der Marel R. P., Franx M., 1993, ApJ, 407, 525

van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000a, AJ, 119, 2038

van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000b, AJ, 119, 2038

Vázquez J. A., Bridges M., Hobson M. P., Lasenby A. N., 2012a, J. Chem. Phys., 6, 006

Vázquez J. A., Bridges M., Hobson M. P., Lasenby A. N., 2012b, J. Chem. Phys., 9, 020

Vikhlinin A. et al., 2009, ApJ, 692, 1033

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Vikhlinin A. et al., 2009, The Astrophysical Journal, 692, 1060

Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 502, 558

Voges W. et al., 1999, A&A, 349, 389

Voit G. M., 2005, Reviews of Modern Physics, 77, 207

von der Linden A. et al., 2014a, MNRAS, 439, 2

von der Linden A. et al., 2014b, MNRAS, 443, 1973

Walker A. G., 1937, Proceedings of the London Mathematical Society, (Series 2) volume 42, p. 90-127, 42, 90

Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Wyn Evans N., Gilmore G., 2009, ApJ, 704, 1274

Wambsganss J., Bode P., Ostriker J. P., 2005, ApJ, 635, L1

Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52

Wetzel A. R., 2011, MNRAS, 412, 49

White S. D. M., 1977, MNRAS, 179, 33

White S. D. M., 1986, The evolution of large-scale structure, Kolb E. W., Turner M. S., Lindley D., Olive K., Seckel D., eds., pp. 228–245

White S. D. M., Efstathiou G., Frenk C. S., 1993a, MNRAS, 262, 1023

White S. D. M., Efstathiou G., Frenk C. S., 1993b, MNRAS, 262, 1023

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993c, Nature, 366, 429

Wilson G. et al., 2009, ApJ, 698, 1943

Wojtak R., Łokas E. L., 2010, MNRAS, 408, 2442

Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., 2009, MNRAS, 399, 812

Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Klypin A., Hoffman Y., 2008, MNRAS, 388, 815

Wolf J., Martinez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R. R., Simon J. D., Avedo F. F., 2010, MNRAS, 406, 1220

Yahil A., Vidal N. V., 1977, ApJ, 214, 347

Yee H. K. C., López-Cruz O., 1999, AJ, 117, 1985

Zabludoff A. I., Franx M., Geller M. J., 1993, ApJ, 419, 47

Zenteno A. et al., 2016, MNRAS, 462, 830

Zhao D. H., Mo H. J., Jing Y. P., Börner G., 2003, MNRAS, 339, 12

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

Zwicky F., Humason M. L., 1964, ApJ, 139, 269

Acknowledgments

Thanks

To Joe, for giving me the opportunity to embark on this journey, and begin my academic career. Thank you for what you taught me, for all the scientific discussions, your guidance, and for making it possible for me to present my work in many conferences and meetings.

To Alex. To you goes my special gratitude. Thank you for your infinite patience with all my questions and doubts. For your time, your help, and your support. I learned a lot from you. And thank you for all the fun times, too!

To the office mates I had during these years. To Nikhel, I-non, Sebastian B., Christina, Corvin, Peter, Sebastian G., Antonio, Luiz, and Maria. Thank you for all the discussions, lunches, dinners, movies, drinks, barbecues, coffee breaks, amazing travels, encouraging words, and archery lessons! To my adoptive office mates, Giorgia, Tea, Steffen and Korbinian, for offering me a warm shelter in cloudy days. To all my USM colleagues, Veronica, Maurilio, Matthias, Joerg, Martin, Thomas, Tadziu, Nico, Ben, Kerstin, Jochen. To Uta, for always landing me a helping hand. My PhD years would have been so much harder without you all!

To the friends I fortunately met in Munich. To Matteo, Aoife, Philip, Malin, Leila, Suhail, Darshan, Damien, Ricard, and Linda. A special thanks goes to Vlas and Manu, for always being there for me, sharing a meal, a drink, a laughter, a frown.

To my "Bologna" people. To Faby, Raffa, Ele, Marco, Mario, Sara, Robb, Marina. Thank you for making me feel like I never really left. To Mario, for always welcoming me with cakes and Japanese food. To Davide, for trusting me with his couch, making all my visits possible, and always cheering me up.

To my parents, my siblings and my extended family, for believing in me. Thank you for your constant support, and bearing with me during my long absences. Thanks for all the trips you made to visit me, and for providing me with joyful times away from duties.

To Steffen, for everything. Thank you for always being by my side, and never giving up on me. Your encouragement made this last steps much more bearable. Thank you for the (countless) inspiring scientific conversations and Skype calls. Thank you for making the kilometers between us feel like just a number. Thank you for being my home, wherever you are.

Thank you.

Let the next adventure begin!